Estimating Multidimensional Density Functions using the Malliavin-Thalmaier Formula

A. Kohatsu-Higa, Kazuhiro Yasuda

July 17, 2008

Abstract: The Malliavin-Thalmaier formula was introduced in [8] as an alternative expression for the density of a multivariate smooth random variable in Wiener space. In comparison with classical integration by parts formulae, this alternative formulation requires the application of the integration by parts formula only once to obtain an expression that can be simulated. Therefore this expression is free from the curse of dimensionality. Unfortunately, when this formula is applied directly in computer simulation, it exhibits unstable behavior. We propose an approximation to the Malliavin-Thalmaier formula in the spirit of the theory of kernel density estimation to solve this problem. In the first part of this paper, we obtain a central limit theorem for the estimation error. And in the latter part, we apply the Malliavin-Thalmaier formula for the calculation of Greeks in finance.

Keywords: Malliavin-Thalmaier formula, Multidimensional density function, Greeks

1 Introduction

Let $(\Omega, \mathcal{F}, P)$ denote a complete probability space carrying a k-dimensional Wiener process W and let $F: \Omega \to \mathbb{R}^d$, $F = (F_1, ..., F_d)$, $d \geq 2$ be a random vector defined in this space. The goal of the present article is to discuss how to simulate the probability density function of F for $d \geq 2$ using Malliavin Calculus. Applications of this problem can be found in a variety of fields where fundamental solutions or density functions can not be explicitely obtained. This problem has attracted some interest due to its financial applications although we frame it here as a general density estimation problem.

The classical integration by parts formula of Malliavin Calculus is an approach that has been suggested by Fournié et. al. [5]. For definitions and results on Malliavin Calculus, we refer the reader to Section 2 of this article where a brief introduction is given or Nualart [9], Theorem 2.1.4 and Proposition 2.1.5 p.102-103 or Sanz-Solé [10], Proposition 5.4 p.67.

Our starting point is an expression for the density of a smooth d-dimensional random vector F. This basic result can be stated as follows.

Let $F = (F_1, ..., F_d)$ be a nondegenerate random vector and G a smooth random variable. We denote by $p_{F,G} = E[G/F = x]p_{F,1}(x)$, where $p_{F,1}(x)$ denotes the density of F. Then there exists a random variable $H_{(1,2,...,d)}(F; 1) \in L^p(\Omega)$ for any $p > 2$ such that

$$p_{F,G}(\hat{\mathbf{x}}) = E\left[\prod_{i=1}^{d} \mathbf{1}_{[0,\infty)}(F_i - \hat{x}_i) H_{(1,2,...,d)}(F; G)\right], \tag{1.1}$$

where $\mathbf{1}_{[0,\infty)}(x)$ denotes the indicator function. Here, for $i = 2, ..., d$,

$$H_{(1)}(F;G) := \sum_{j=1}^{d} \delta\Big(G(\gamma_F^{-1})^{1j} DF_j\Big),$$
$$H_{(1,...,i)}(F;G) := \sum_{j=1}^{d} \delta\Big(H_{(1,...,i-1)}(F;G)(\gamma_F^{-1})^{ij} DF_j\Big). \tag{1.2}$$

Here δ denotes the adjoint operator of the Malliavin derivative operator D and γ_F the Malliavin covariance matrix of F.

In particular, we remark that δ is an extension of the Itô integral that also integrates non-adapted processes and is usually called the Skorohod integral. The definition of $H_{(1,...,i)}(F;1)$ in iterative form in (1.2), shows that in order to compute this expression one requires the calculation of i-iterated stochastic integrals.

The Skorohod integral being a non-adapted integral is not easy to simulate in iterative form and therefore the above expression takes a relatively large amount of time to be simulated when d is big unless an explicit simple expression for $H_{(1,...,d)}(F;G)$ is obtained. Besides this problem, one also encounters problems of high variance and therefore variance reduction methods have to be incorporated making the problem even less tractable from an applied point of view.

Recently, Malliavin and Thalmaier [8] (Section 4.5.) introduced an alternative integration by parts formula that seems to alleviate the computational burden for simulation of densities in high dimension. In fact, Malliavin and Thalmaier express the multi-dimensional delta function as

$$\delta_0(\mathbf{x}) = \Delta Q_d(\mathbf{x}),$$

where $\Delta = \sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2}$ is the Laplace operator and Q_d is the fundamental solution of the Poisson equation. Then for $\hat{\mathbf{x}} \in \mathbb{R}^d$, they obtain the following representation for the density of F

$$p_{F,G}(\hat{\mathbf{x}}) = E\left[\sum_{i=1}^{d} \frac{\partial}{\partial x_i} Q_d(F - \hat{\mathbf{x}}) H_{(i)}(F;G)\right]. \tag{1.3}$$

Therefore one needs to simulate $H_{(i)}(F;G)$ which involves only one Skorohod integral instead of the previous d-iterated Skorohod integrals in (1.2).

In fact, if we partition the time interval in N intervals in order to carry out simulations of the increments of the Wiener process, then the iterated Skorohod integrals appearing in (1.1) will require the calculation over N^d cross-intervals. Instead formula (1.3) only requires Nd.

In principle, one expects then that the calculation time will be highly reduced. Nevertheless, the high variance problem in formula (1.1) is taken to an extreme as the variance of the estimator in (1.3) is infinite. This problem appears because the limit of $\frac{\partial}{\partial x_i} Q_d(\mathbf{x})$ at $\mathbf{x} = 0$ is ∞, although the expectation in (1.3) is finite.

Therefore we propose a slightly modified estimator that depends on a modification parameter h which will converge to the function $\frac{\partial}{\partial x_i} Q_d(\mathbf{x})$ as $h \to 0$. This will generate a small bias and a large variance which is not infinite. Then we control the explosive behavior of the variance using the number of simulations. This type of calculation is common in kernel density estimation (KDE) methods where this technique has been very effective. The main difference between traditional KDE theory and the proposal in this paper is that although the modification we propose here is mathematically natural it does not correspond to any of the classical estimation methods studied in KDE theory.

In order to "tune the parameter h" (expression used in KDE, meaning how to choose h) we obtain in Section 3 the bias of the estimation procedure. In Section 4 we study the $L^2(\Omega)$ error of estimation to finally obtain in Section 5, the central limit theorem that shows how to tune the parameters of the estimation procedure. In Section 6 we apply the Malliavin-Thalmaier formula to finance, especially to the calculation of Greeks, in the spirit of Fournié *et al* [5] where the one dimensional examples are considered. We give an expression for Greeks using the Malliavin-Thalmaier formula. In particular, the weights are free from the curse of dimensionality. That is, the expression does not have a d-iterated Skorohod integral.

The article closes in section 7, with the discussion of various simulation results. In particular, we concentrate on the simulation of the density of the stock value and volatility in the Heston model.

In order to avoid long proofs we have moved to an Appendix all technical details leaving in the proofs of the main theorems the essential ideas.

Also note that the expression in (1.1) corresponds to a density only in the case that $G = 1$. To avoid introducing further terminology, we will keep referring to $p_{F,G}(\hat{\mathbf{x}})$ as the "density".

2 Preliminaries

Let us introduce some notations and basic definitions of Malliavin Calculus. For a multi-index $\alpha = (\alpha_1, ..., \alpha_m) \in \{1, ..., d\}^m$, we denote by $|\alpha| = m$ the length of the multi-index.

2.1 Malliavin Calculus

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space. Let W be a k-dimensional Wiener process on the time interval $[0, T]$.

We denote by $C_p^\infty(\mathbb{R}^n)$ the set of all infinitely differentiable functions $f : \mathbb{R}^n \to \mathbb{R}$ such that f and all of its partial derivatives have at most polynomial growth.

Let $\mathcal{S}$ denote the class of *smooth* random variables of the form

$$F = f(W(t_1), ..., W(t_n)), \tag{2.1}$$

where $f \in C_p^\infty(\mathbb{R}^n)$, $t_1, ..., t_n \in [0, T]$, and $n \geq 1$.

If F has the form (2.1) we define its derivative $D_s^i F$, $i = 1, ..., k$ as

$$D_s^i F = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(W(t_1), ..., W(t_n)) 1_{[0,t_i]}(s).$$

We will denote the domain of D in $L^p(\Omega)$ by $\mathbb{D}^{1,p}$. This space is the closure of the class of smooth random variables $\mathcal{S}$ with respect to the norm

$$\|F\|_{1,p} = \left\{ E\Big[|F|^p\Big] + E\Big[\|DF\|^p_{L^2[0,T]}\Big] \right\}^{\frac{1}{p}}.$$

We can define the iteration of the operator D in such a way that for a smooth random variable F, the derivative $D^k F$ is a random variable with values on $L^2[0, T]^{\otimes k}$. Then for every $p \geq 1$ and $k \in \mathbb{N}$ we introduce a seminorm on $\mathcal{S}$ defined by

$$\|F\|^p_{n,p} = E\Big[|F|^p\Big] + \sum_{j=1}^{n} E\Big[\|D^j F\|^p_{L^2[0,T]^{\otimes j}}\Big].$$

For any real $p \geq 1$ and any natural number $n \geq 0$, we will denote by $\mathbb{D}^{n,p}$ the completion of the family of smooth random variables $\mathcal{S}$ with respect to the norm $\|\cdot\|_{n,p}$. Note that $\mathbb{D}^{j,p} \subset \mathbb{D}^{n,q}$ if $j \geq n$ and $p \geq q$.

Consider the intersection

$$\mathbb{D}^\infty = \bigcap_{p\geq 1} \bigcap_{n \geq 1} \mathbb{D}^{n,p}.$$

Then $\mathbb{D}^\infty$ is a complete, countably normed, metric space.

We will denote by δ *the adjoint of the operator* D as an unbounded operator from $L^2(\Omega)$ into $L^2(\Omega; L^2[0,T])$. That is, the domain of δ, denoted by $\mathrm{Dom}(\delta)$, is the set of $L^2[0,T]$-valued square integrable random variables u such that

$$|E[< DF, u >_{L^2[0,T]}]| \leq c\|F\|_2,$$

for all $F \in \mathbb{D}^{1,2}$, where c is some positive constant depending on u. (here $\|\cdot\|_2$ denotes the $L^2(\Omega)$-norm.) We remark here that any L^2 integrable adapted process u, belongs to the domain of δ. Furthermore one can prove that in such a case $\delta(u)$ is the Itô integral of u. In general, $\delta(u)$ is called the Skorohod integral of u. A property of δ that is frequently used is: Let $G \in \mathbb{D}^{1,2}$ be a real-valued random variable such that $Gu \in L^2(\Omega, L^2[0,T]))$, then

$$\delta(Gu) = G\delta(u) - < DG, u >_{L^2[0,T]} \tag{2.2}$$

where we suppose that the right hand side is integrable.

Suppose that $F = (F_1, ..., F_d)$ is a random vector whose components belong to the space $\mathbb{D}^{1,1}$. We associate with F the following random symmetric nonnegative definite matrix:

$$\gamma_F = \Big(< DF_i, DF_j >_{L^2[0,T]} \Big)_{1 \leq i,j \leq d}.$$

This matrix is called the *Malliavin covariance matrix* of the random vector F.

Definition 2.1 *We say that the random vector $F = (F_1, ..., F_d) \in (\mathbb{D}^\infty)^d$ is nondegenerate if the matrix γ_F is invertible a.s. and*

$$(\det \gamma_F)^{-1} \in \bigcap_{p \geq 1} L^p(\Omega).$$

It is well known that if F is nondegenerate and $G \in \mathbb{D}^\infty$ then $p_{F,G}$ exists and is smooth and in particular one obtains (1.1).

2.2 Malliavin-Thalmaier Representation of Multi-Dimensional Density Functions

We represent the delta function by

$$\delta_{\mathbf{0}}(\mathbf{x}) = \Delta Q_d(\mathbf{x}) \quad \text{for } \mathbf{x} \in \mathbb{R}^d,\ d \geq 2,$$

in the following sense (see Evans [4], p.25). If f is a twice continuously differentiable function with compact support, then the solution of the Poisson equation $\Delta u = f$ is given by the convolution $Q_d * f$ where the fundamental solution (also called Poisson kernel) Q_d has the following explicit form;

$$Q_2(\mathbf{x}) := a_2^{-1} \ln |\mathbf{x}| \quad \text{and} \quad Q_d(\mathbf{x}) := -a_d^{-1} \frac{1}{|\mathbf{x}|^{d-2}} \quad \text{for } d \geq 3.$$

Here a_d is the area of the unit sphere in $\mathbb{R}^d$. The derivative of the Poisson kernel is

$$\frac{\partial}{\partial x_i} Q_d(\mathbf{x}) = A_d \frac{x_i}{|\mathbf{x}|^d},$$

where $i = 1, ..., d$, $A_2 := a_2^{-1}$ and for $d \geq 3$, $A_d := a_d^{-1}(d-2)$.

Definition 2.1 *Given the $\mathbb{R}^d$-valued random vector F and the $\mathbb{R}$-valued random variable G, a multi-index α and a power $p \geq 1$ we say that there is an integration by parts formula (IBP formula) if there exists a random variable $H_\alpha(F;G) \in L^p(\Omega)$ such that*

$$IP_{\alpha,p}(F,G): \quad E\left[\frac{\partial^{|\alpha|}}{\partial \mathbf{x}^\alpha} f(F) G\right] = E\Big[f(F) H_\alpha(F;G)\Big] \quad \text{for all } f \in C_0^{|\alpha|}(\mathbb{R}^d).$$

Related to the Malliavin-Thalmaier formula, Bally and Caramellino [2], have obtained the following result, which gives specific conditions for the Malliavin-Thalmaier formula to hold.

Proposition 2.1 *(Bally, Caramellino [2]) Suppose that for some $p > 1$*

$$\sup_{|\mathbf{a}|\le R} E\left[\left|\frac{\partial}{\partial x_i}Q_d(F-\mathbf{a})\right|^{\frac{p}{p-1}} + \left|Q_d(F-\mathbf{a})\right|^{\frac{p}{p-1}}\right] < \infty \quad \text{for all } R>0,\ \mathbf{a}\in\mathbb{R}^d. \tag{2.3}$$

Then for $\hat{\mathbf{x}} \in \mathbb{R}^d$, we have,
(i). If $IP_{i,p}(F;G)$, $i = 1, ..., d$, holds then the law of F is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$ and the density $p_{F,G}$ is represented as

$$p_{F,G}(\hat{\mathbf{x}}) = E\left[\sum_{i=1}^{d} \frac{\partial}{\partial x_i} Q_d(F-\hat{\mathbf{x}})H_{(i)}(F;G)\right]. \tag{2.4}$$

(ii). If $IP_{\alpha,p}(F;G)$ holds for every multi-index α with $|\alpha| \le m+1$ then $p_{F,G} \in C^m(\mathbb{R}^d)$ and for every multi-index ρ with $|\rho| \le m$ one has

$$\frac{\partial^{|\rho|}}{\partial \mathbf{x}^\rho} p_{F,G}(\hat{\mathbf{x}}) = E\left[\sum_{i=1}^{d} \frac{\partial}{\partial x_i} Q_d(F-\hat{\mathbf{x}})H_{(i,\rho)}(F;G)\right].$$

The heuristic idea of the above proof is to use the IBP formula as follows

$$p_{F,G}(\hat{\mathbf{x}}) = E\Big[\Delta Q_d(F-\hat{\mathbf{x}})G\Big] = \sum_{i=1}^{d} E\left[\frac{\partial^2}{\partial x_i^2}Q_d(F-\hat{\mathbf{x}})G\right] = E\left[\sum_{i=1}^{d}\frac{\partial}{\partial x_i}Q_d(F-\hat{\mathbf{x}})H_{(i)}(F;G)\right].$$

Next we impose conditions to assure that the assumptions of Proposition 2.1 are satisfied. The proof is given in the Appendix.

Corollary 2.1 *If $F = (F_1, ..., F_d)$ is a nondegenerate random vector and $G \in \mathbb{D}^\infty$, then the probability density function of the random vector F is, for $\hat{\mathbf{x}} \in \mathbb{R}^d$,*

$$p_{F,G}(\hat{\mathbf{x}}) = E\left[\sum_{i=1}^{d} \frac{\partial}{\partial x_i} Q_d(F-\hat{\mathbf{x}})H_{(i)}(F;G)\right].$$

Assumption 2.1 *From now on, we always assume that $F = (F_1, ..., F_d)$ is a d-dimensional nondegenerate random variable and $G \in \mathbb{D}^\infty$.*

3 Bias Error Estimation

In this section, we find the rate of convergence of the modified estimator of the density at $\hat{\mathbf{x}} \in \mathbb{R}^d$. From Assumption 2.1, $IP_{\alpha,p}(F;G)$ will always hold (see Nualart [9], Proposition 2.1.4, p.100 or Sanz-Solé [10], Proposition 5.4 p.67). We start with some definitions and notations.

Definitions and Notations

1. For $h > 0$ and $\mathbf{x} \in \mathbb{R}^d$, define $|\cdot|_h$ by

$$|\mathbf{x}|_h := \sqrt{\sum_{i=1}^{d} x_i^2 + h}.$$

Without loss of generality, we assume $0 < h < 1$.

2. For $i = 1, ..., d$, define the following approximation to Q_d, for $\mathbf{x} \in \mathbb{R}^d$,

$$Q_d^h(\mathbf{x}) = \begin{cases} a_2^{-1} \ln |\mathbf{x}|_h & ; d = 2 \\ -a_d^{-1} \frac{1}{|\mathbf{x}|_h^{d-2}} & ; d \ge 3 \end{cases}$$

Then we have that
$$\frac{\partial}{\partial x_i} Q_d^h(\mathbf{x}) = A_d \frac{x_i}{|\mathbf{x}|_h^d}.$$

3. Then we define the approximation to the density function of F, for $\mathbf{x} \in \mathbb{R}^d$, as
$$p_{F,G}^h(\mathbf{x}) := E\left[\sum_{i=1}^{d} \frac{\partial}{\partial x_i} Q_d^h(F-\mathbf{x}) H_{(i)}(F;G)\right]. \tag{3.1}$$

4. Consider a function η which satisfies;
$$\begin{cases} (i). & \eta \in C_0^\infty(\mathbb{R}^d), \quad \eta(\mathbf{x}) \geq 0 \quad (\mathbf{x} \in \mathbb{R}^d), \\ (ii). & \operatorname{supp}(\eta) \subset \{\mathbf{x} \in \mathbb{R}^d \mid |\mathbf{x}| \leq 1\}, \\ (iii). & \int_{\mathbb{R}^d} \eta(\mathbf{x}) d\mathbf{x} = 1, \\ (iv). & \eta(\mathbf{x}) \text{ is symmetric, that is, } \eta(\mathbf{x}) = \eta(\mathbf{y}) \text{ when } |\mathbf{x}| = |\mathbf{y}| \text{ for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^d. \end{cases}$$

5. For each $\varepsilon > 0$, we define $\eta_\varepsilon(\mathbf{x})$ as
$$\eta_\varepsilon(\mathbf{x}) := \frac{1}{\varepsilon^d} \eta\left(\frac{\mathbf{x}}{\varepsilon}\right).$$

6. We define $\tilde{\eta}_\varepsilon(\mathbf{x})$;
$$\tilde{\eta}_\varepsilon(\mathbf{x}) := \int_{-\infty}^{x_d} \cdots \int_{-\infty}^{x_1} \eta_\varepsilon(\mathbf{y}) dy_1 \ldots dy_d. \ (\leq 1 \quad \text{from } \mathbf{4.})$$

7. We often use the spherical coordinates. To avoid long expressions we define $\Theta := (\Theta_1, ..., \Theta_d)^*$ as the coordinate change
$$\begin{aligned} r\Theta_1 :=& r\cos(\theta_1)\cos(\theta_2)\cdots\cos(\theta_{d-2})\cos(\theta_{d-1}), \\ r\Theta_i :=& r\cos(\theta_1)\cdots\cos(\theta_{d-i})\sin(\theta_{d-i+1}) \quad \text{for } i = 2, ..., d, \end{aligned}$$
where $0 \leq r < \infty$, $-\frac{\pi}{2} \leq \theta_i \leq \frac{\pi}{2}$, $i = 1, ..., d-2$, $0 \leq \theta_{d-1} \leq 2\pi$. Set $s_i = \sin\theta_i$, $c_i = \cos\theta_i$ for $i = 1, ..., d-1$.

We will give some preparatory Lemmas for the following section.

Lemma 3.1 *For $m \in \mathbb{N} \cup \{0\}$, let $\alpha \in \{1, ..., d\}^m$, be any multi-index. Then for $\hat{\mathbf{x}} = (\hat{x}_1, ..., \hat{x}_d) \in \mathbb{R}^d$, there exists some constant C such that for $p \geq 1$,*
$$\lim_{\varepsilon \to 0} \left| \frac{\partial^{|\alpha|}}{\partial \mathbf{x}^\alpha} E\left[\eta_\varepsilon(F-\hat{\mathbf{x}})G\right] \right| \leq \frac{C}{1+|\hat{\mathbf{x}}|^p}.$$

Proof. It is enough to consider the case $p \in \mathbb{N}$. In such a case, we have
$$\begin{aligned} \left| \lim_{\varepsilon \to 0} (1+|\hat{\mathbf{x}}|^p) \frac{\partial^m}{\partial \mathbf{x}^\alpha} E\left[\eta_\varepsilon(F-\hat{\mathbf{x}})G\right] \right| &= \left| \lim_{\varepsilon \to 0} (1+|\hat{\mathbf{x}}|^p) E\left[\eta_\varepsilon(F-\hat{\mathbf{x}}) H_\alpha(F,G)\right] \right| \\ &\leq \lim_{\varepsilon \to 0} \left| E\left[\eta_\varepsilon(F-\hat{\mathbf{x}})(1+(|F|+\varepsilon)^p) H_\alpha(F,G)\right] \right| \\ &\leq C_p E\left[\left| H_{(1,\ldots,d)}(F;(1+|F|^{2p}) H_\alpha(F,G)) \right|\right] < +\infty. \end{aligned}$$
■

Lemma 3.2 *The following holds, for $\hat{\mathbf{x}} = (\hat{x}_1, ..., \hat{x}_d) \in \mathbb{R}^d$,*
$$\lim_{\varepsilon \to 0} E\Big[\eta_\varepsilon(F-\hat{\mathbf{x}})G\Big] = E[G|F=\hat{\mathbf{x}}] p_{F,1}(\hat{\mathbf{x}}).$$

Proof. Set $z_i = F_i - x_i$ $(i = 1, ..., d)$. By the dominated convergence theorem and the properties of η_ε, stated in **4.**, and Fubini's theorem, we have for $\varphi \in C_0^\infty(\mathbb{R}^d)$,

$$\begin{aligned}\int_{\mathbb{R}^d}\left(\lim_{\varepsilon\to 0} E\left[\eta_\varepsilon(F-\hat{\mathbf{x}})G\right]\right)\varphi(\hat{\mathbf{x}})d\hat{\mathbf{x}} &= \lim_{\varepsilon\to 0} E\left[\int_{\mathbb{R}^d}\eta_\varepsilon(\mathbf{z})\varphi(F+\mathbf{z})d\mathbf{z}G\right]\\ &= E\Big[\varphi(F)G\Big]\\ &= \int_{\mathbb{R}^d}\varphi(\hat{\mathbf{x}})E\Big[G\Big|F=\hat{\mathbf{x}}\Big]p_{F,1}(\hat{\mathbf{x}})d\hat{\mathbf{x}}.\end{aligned}$$

■

The next result gives the order of the error of the approximation to the density.

Theorem 3.1 *Let F be a nondegenerate random vector and $G \in \mathbb{D}^\infty$, then for $\hat{\mathbf{x}} = (\hat{x}_1, ..., \hat{x}_d) \in \mathbb{R}^d$,*

$$p_{F,G}(\hat{\mathbf{x}}) - p^h_{F,G}(\hat{\mathbf{x}}) = C_1^{\hat{\mathbf{x}}} h \ln\frac{1}{h} + C_2^{\hat{\mathbf{x}}} h + o(h),$$

where

$$C_1^{\hat{\mathbf{x}}} := \sum_{i=1}^d C_{1,i}^{\hat{\mathbf{x}}} \quad \text{and} \quad C_2^{\hat{\mathbf{x}}} := \sum_{i=1}^d \left\{ C_{2,i}^{\hat{\mathbf{x}}} + \sum_{j,k=1}^d C_{3,i,j,k}^{\hat{\mathbf{x}}} + C_{4,i}^{\hat{\mathbf{x}}} \right\},$$

and the constants appearing above are defined in Lemmas 8.3, 8.4 and 8.5 in the Appendix.

Proof. As we will have to change from rectangular to spherical coordinates, set $y_1 - \hat{x}_1 = r\Theta_1$ and $y_i - \hat{x}_i = r\Theta_i$ for $i = 2, ..., d$.

By using Lemma 3.2 and spherical coordinates,

$$\begin{aligned}p_{F,G}(\hat{\mathbf{x}}) - p^h_{F,G}(\hat{\mathbf{x}}) &= E\left[\sum_{i=1}^d\left(\frac{\partial}{\partial x_i}Q_d(F-\hat{\mathbf{x}}) - \frac{\partial}{\partial x_i}Q_d^h(F-\hat{\mathbf{x}})\right)H_{(i)}(F;G)\right]\\ &= A_d\sum_{i=1}^d\int_{\mathbb{R}^d}\left(\frac{y_i-\hat{x}_i}{|\mathbf{y}-\hat{\mathbf{x}}|^d} - \frac{y_i-\hat{x}_i}{|\mathbf{y}-\hat{\mathbf{x}}|_h^d}\right)\left(\lim_{\varepsilon\to\infty}E\Big[\eta_\varepsilon(F-\mathbf{y})H_{(i)}(F;G)\Big]\right)dy_1\cdots dy_d\\ &= A_d\sum_{i=1}^d\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\left(\int_0^1+\int_1^\infty\right)\frac{(r^2+h)^{\frac{d}{2}}-r^d}{(r^2+h)^{\frac{d}{2}}}\Theta_i c_1^{d-2}\cdots c_{d-2}\left(\lim_{\varepsilon\to 0}\Phi^F_{i,\varepsilon}(r\Theta+\hat{\mathbf{x}})\right)drd\theta_1...d\theta_{d-1},\end{aligned}$$

where $\Phi^F_{i,\varepsilon}(\mathbf{y}) := E\left[\eta_\varepsilon(F-\mathbf{y})H_{(i)}(F;G)\right]$ for $i = 1, ..., d$. Here note that the limits appearing in the above formula exist due to Lemma 3.1 and Lemma 3.2.

Next, we consider the integral for $r \in [0, 1]$ where the following Taylor formula is used

$$\Phi^F_{i,\varepsilon}(r\Theta+\hat{\mathbf{x}}) = \Phi^F_{i,\varepsilon}(\hat{\mathbf{x}}) + \sum_{j=1}^d r\Theta_j\frac{\partial}{\partial y_j}\Phi^F_{i,\varepsilon}(\hat{\mathbf{x}}) + \frac{1}{2}\sum_{j,k=1}^d r^2\Theta_j\Theta_k\int_0^1\frac{\partial^2}{\partial y_k\partial y_j}\Phi^F_{i,\varepsilon}(\hat{\mathbf{x}}+\gamma r\Theta)d\gamma.$$

This leads to three terms, whose orders of convergence are analyzed respectively in Lemmas 8.2, 8.3 and 8.4 in the Appendix. Finally, the integral term for $r \in [1, +\infty)$ is analyzed in Lemma 8.5 in the Appendix. Therefore one obtains that

$$p_{F,G}(\hat{\mathbf{x}}) - p^h_{F,G}(\hat{\mathbf{x}}) = \sum_{i=1}^d\left\{C_{1,i}^{\hat{\mathbf{x}}}h\ln\frac{1}{h} + C_{2,i}^{\hat{\mathbf{x}}}h + o(h) + \sum_{j,k=1}^d C_{3,i,j,k}^{\hat{\mathbf{x}}}h + o(h) + C_{4,i}^{\hat{\mathbf{x}}}h + o(h)\right\}.$$

The constants are explicitly given in the Appendix.

■

4 Estimation of the L^2-error of the Approximation

In this section, we compute the rate at which the L^2-error of the estimator diverges. That is,

$$E\left[\left(\sum_{i=1}^{d}\frac{\partial}{\partial x_i}Q_d^h(F-\hat{\mathbf{x}})H_{(i)}(F;G)-p_{F,G}(\hat{\mathbf{x}})\right)^2\right]$$
$$=E\left[\left(\sum_{i=1}^{d}\frac{\partial}{\partial x_i}Q_d^h(F-\hat{\mathbf{x}})H_{(i)}(F;G)\right)^2\right]+2p_{F,G}(\hat{\mathbf{x}})\Big\{p_{F,G}(\hat{\mathbf{x}})-p_{F,G}^h(\hat{\mathbf{x}})\Big\}-p_{F,G}(\hat{\mathbf{x}})^2. \quad (4.1)$$

Therefore it is enough to estimate the rate of divergence of the first term in (4.1) as the second term converges to 0 (proved in Theorem 3.1) and the third is a constant. The term we will calculate is then

$$E\left[\left(\sum_{i=1}^{d}\frac{\partial}{\partial x_i}Q_d^h(F-\hat{\mathbf{x}})H_{(i)}(F;G)\right)^2\right]=\sum_{i,j=1}^{d}E\left[\frac{\partial}{\partial x_i}Q_d^h(F-\hat{\mathbf{x}})\frac{\partial}{\partial x_j}Q_d^h(F-\hat{\mathbf{x}})H_{(i)}(F;G)H_{(j)}(F;G)\right].$$

Let $\hat{\Phi}^F_{i,j,\varepsilon}(\mathbf{y}):=E\left[\eta_\varepsilon(F-\mathbf{y})H_{(i)}(F;G)H_{(j)}(F;G)\right]$ for $i,\ j=1,...,d$.

4.1 Case $d=2$

Theorem 4.1 *Let F be a nondegenerate random vector and $G\in\mathbb{D}^\infty$. Then for $d=2$,*

$$E\left[\left(\sum_{i=1}^{2}\frac{\partial}{\partial x_i}Q_2^h(F-\hat{\mathbf{x}})H_{(i)}(F;G)-p_{F,G}(\hat{\mathbf{x}})\right)^2\right]=C_3^{\hat{\mathbf{x}}}\ln\frac{1}{h}+O(1)\quad(\hat{\mathbf{x}}\in\mathbb{R}^d),$$

where $C_3^{\hat{\mathbf{x}}}:=\sum_{i=1}^2 C_{5,i}^{\hat{\mathbf{x}}}$ and the constants $C_{5,i}^{\hat{\mathbf{x}}}$ are defined in Lemma 8.6 in the Appendix.

Proof. Set $y_i-\hat{x}_i=r\Theta_i$ for $i=1,2$. For $i,j=1,2$, by using Lemma 3.2, Taylor expansion and spherical coordinates,

$$E\left[\frac{\partial}{\partial x_i}Q_2^h(F-\hat{\mathbf{x}})\frac{\partial}{\partial x_j}Q_2^h(F-\hat{\mathbf{x}})H_{(i)}(F;G)H_{(j)}(F;G)\right]$$
$$=A_2^2\int_{\mathbb{R}^2}\frac{(y_i-\hat{x}_i)(y_j-\hat{x}_j)}{|\mathbf{y}-\hat{\mathbf{x}}|_h^4}\left(\lim_{\varepsilon\to 0}\hat{\Phi}^F_{i,j,\varepsilon}(\mathbf{y})\right)dy_1dy_2$$
$$=A_2^2\int_0^{2\pi}\int_0^{2|\hat{\mathbf{x}}|+1}r\frac{r^2\Theta_i\Theta_j}{(r^2+h)^2}\left\{\lim_{\varepsilon\to 0}\left(\hat{\Phi}^F_{i,j,\varepsilon}(\hat{\mathbf{x}})+\sum_{k=1}^{2}r\Theta_k\int_0^1\frac{\partial}{\partial y_k}\hat{\Phi}^F_{i,j,\varepsilon}(\hat{\mathbf{x}}+\gamma r\Theta)d\gamma\right)\right\}drd\theta$$
$$+A_2^2\int_0^{2\pi}\int_{2|\hat{\mathbf{x}}|+1}^{\infty}r\frac{r^2\Theta_i\Theta_j}{(r^2+h)^2}\left(\lim_{\varepsilon\to 0}\hat{\Phi}^F_{i,j,\varepsilon}(r\Theta+\hat{\mathbf{x}})\right)drd\theta. \quad (4.2)$$

Then by using Lemma 8.6, Lemma 8.7 and Lemma 8.8, we obtain

$$(4.2)=\sum_{i=1}^{2}C_{5,i}^{\hat{\mathbf{x}}}\ln\frac{1}{h}+O(1).$$

■

4.2 Case $d\geq 3$

Theorem 4.2 *Let F be a nondenegerate random vector and $G\in\mathbb{D}^\infty$. Then for $d\geq 3$,*

$$E\left[\left(\sum_{i=1}^{d}\frac{\partial}{\partial x_i}Q_d^h(F-\hat{\mathbf{x}})H_{(i)}(F;G)-p_{F,G}(\hat{\mathbf{x}})\right)^2\right]=C_4^{\hat{\mathbf{x}}}\frac{1}{h^{\frac{d}{2}-1}}+o\left(\frac{1}{h^{\frac{d}{2}-1}}\right)\quad(\hat{\mathbf{x}}\in\mathbb{R}^d),$$

where $C_4^{\hat{\mathbf{x}}}=\sum_{i=1}^d C_{8,i}^{\hat{\mathbf{x}}}$ and the constants $C_{8,i}^{\hat{\mathbf{x}}}$ are defined in Lemma 8.10.

Proof. Let $y_i - \hat{x}_i = r\Theta_i$ for $i = 1, ..., d$. For $i, j = 1, ..., d$, by using Lemma 3.2, Taylor expansion and spherical coordinates,

$$
\begin{aligned}
&E\left[\frac{\partial}{\partial x_i} Q_d^h(F - \hat{\mathbf{x}}) \frac{\partial}{\partial x_j} Q_d^h(F - \hat{\mathbf{x}}) H_{(i)}(F; G) H_{(j)}(F; G)\right] \\
&= A_d^2 \int_{\mathbb{R}^d} \frac{(y_i - \hat{x}_i)(y_j - \hat{x}_j)}{|\mathbf{y} - \hat{\mathbf{x}}|_h^{2d}} \left(\lim_{\varepsilon \to 0} \hat{\Phi}^F_{i,j,\varepsilon}(\mathbf{y})\right) dy_1 ... dy_d \\
&= A_d^2 \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cdots \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^1 \frac{r^2 \Theta_i \Theta_j}{(r^2 + h)^d} r^{d-1} c_1^{d-2} \cdots c_{d-2} \qquad (4.3) \\
&\qquad \times \left\{ \lim_{\varepsilon \to 0} \left(\hat{\Phi}^F_{i,j,\varepsilon}(\hat{\mathbf{x}}) + \sum_{k=1}^d r\Theta_k \int_0^1 \frac{\partial}{\partial y_k} \hat{\Phi}^F_{i,j,\varepsilon}(\hat{\mathbf{x}} + \gamma r \Theta) d\gamma \right) \right\} dr d\theta_1 ... d\theta_{d-1} \\
&\quad + A_d^2 \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cdots \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_1^\infty \frac{r^2 \Theta_i \Theta_j}{(r^2 + h)^d} r^{d-1} c_1^{d-2} \cdots c_{d-2} \left(\lim_{\varepsilon \to 0} \hat{\Phi}^F_{i,j,\varepsilon}(\hat{\mathbf{x}} + r\Theta)\right) dr d\theta_1 ... d\theta_{d-1}.
\end{aligned}
$$

Then from Lemma 8.10, Lemma 8.11 and Lemma 8.12, we can obtain our result. ■

Remark. In particular, for $h = 0$ one obtains that the variance of the Malliavin-Thalmaier estimator is infinite. We also point out that this situation also appears in kernel density estimation theory. In particular, one uses as estimator $h^{-1}K(\frac{F-x}{h})$ where h is the tunning parameter and K is a smooth density function with mean 0 and finite moments. In this case, the bias is of order $O(h^2)$ and the L^2-error is of order $O(\frac{1}{h^{d/2}})$. In that situation, as we will do in the next section, the solution is to use the sample size in order to obtain the convergence of the estimator.

5 The Central Limit Theorem

Obviously when performing simulations, one is also interested in obtaining confidence intervals and therefore the Central Limit Theorem is useful in such a situation. In what follows $\Rightarrow$ denotes weak convergence and the index $j = 1, ..., N$ denotes N independent copies of the respective random variables. The symbol $\lfloor \cdot \rfloor$, denotes the greatest integer function.

Theorem 5.1 *Let Z be a random variable with standard normal distribution and let $(F^{(j)}, G^{(j)}) \in (\mathbb{D}^\infty)^d \times \mathbb{D}^\infty$, $j \in \mathbb{N}$ be a sequence of independent identically distributed random vectors.*
(i). *When $d = 2$, set $n = \left\lfloor \frac{C}{h \ln \frac{1}{h}} \right\rfloor$ and $N = \left\lfloor \frac{C^2}{h^2 \ln \frac{1}{h}} \right\rfloor$ for some positive constant C fixed throughout. Then as $h \to 0$*

$$
n\left(\frac{1}{N}\sum_{j=1}^N \sum_{i=1}^2 \frac{\partial}{\partial x_i} Q_2^h(F^{(j)} - \hat{\mathbf{x}}) H_{(i)}(F; G)^{(j)} - p_{F,G}(\hat{\mathbf{x}})\right) \Longrightarrow \sqrt{C_3^{\hat{\mathbf{x}}}} Z - C_1^{\hat{\mathbf{x}}} C, \qquad (5.1)
$$

where $H_{(i)}(F; G)^{(j)}$, $i = 1, ..., d$, $j = 1, ..., N$, denotes the weight obtained in the j-th independent simulation (the same that generates $F^{(j)}$ and $G^{(j)}$).
(ii). *When $d \geq 3$, set $n = \left\lfloor \frac{C}{h \ln \frac{1}{h}} \right\rfloor$ and $N = \left\lfloor \frac{C^2}{h^{\frac{d}{2}+1} (\ln \frac{1}{h})^2} \right\rfloor$ for some positive constant C fixed throughout. Then as $h \to 0$*

$$
n\left(\frac{1}{N}\sum_{j=1}^N \sum_{i=1}^d \frac{\partial}{\partial x_i} Q_d^h(F^{(j)} - \hat{\mathbf{x}}) H_{(i)}(F; G)^{(j)} - p_{F,G}(\hat{\mathbf{x}})\right) \Longrightarrow \sqrt{C_4^{\hat{\mathbf{x}}}} Z - C_1^{\hat{\mathbf{x}}} C. \qquad (5.2)
$$

Proof. Consider

$$n\left(\frac{1}{N}\sum_{j=1}^{N}\sum_{i=1}^{d}\frac{\partial}{\partial x_i}Q_d^h(F^{(j)}-\hat{\mathbf{x}})H_{(i)}\left(F;G\right)^{(j)}-p_{F,G}(\hat{\mathbf{x}})\right)$$
$$=\frac{n}{N}\sum_{j=1}^{N}\left\{\sum_{i=1}^{d}\frac{\partial}{\partial x_i}Q_d^h(F^{(j)}-\hat{\mathbf{x}})H_{(i)}\left(F;G\right)^{(j)}-p^h_{F,G}(\hat{\mathbf{x}})\right\}+n\Big(p^h_{F,G}(\hat{\mathbf{x}})-p_{F,G}(\hat{\mathbf{x}})\Big).$$

Due to the definition of n and Theorem 3.1 we have that the second term above converges to $-C_1^{\hat{\mathbf{x}}}C$. Therefore it only remains to prove a central limit theorem for $\frac{n}{N}\sum_{j=1}^{N}\zeta_j^h$ where

$$\zeta_j^h:=\sum_{i=1}^{d}\frac{\partial}{\partial x_i}Q_d^h(F^{(j)}-\hat{\mathbf{x}})H_{(i)}\left(F;G\right)^{(j)}-p^h_{F,G}(\hat{\mathbf{x}}).$$

Note that $\{\zeta_j^h\}$ is a sequence of the i.i.d. random variables with $E[\zeta_1^h]=0$.

To prove this, we compute the characteristic function of $\frac{n}{N}\sum_{j=1}^{N}\zeta_j^h$. By Taylor expansion, Lemma 8.13 and Lemma 8.14,

$$E\left[\exp\left(\frac{\sqrt{-1}un}{N}\sum_{j=1}^{N}\zeta_j^h\right)\right]=\left\{1-\frac{1}{N}\left(\frac{1}{2}\frac{u^2n^2}{N}E\left[\left(\zeta_1^h\right)^2\right]+N\times\mathcal{R}\right)\right\}^N\longrightarrow\exp\left(-\frac{u^2}{2}C'_{\hat{\mathbf{x}}}\right),$$

where when $d=2$, $C'_{\hat{\mathbf{x}}}=C_3^{\hat{\mathbf{x}}}$ and when $d\geq 3$, $C'_{\hat{\mathbf{x}}}=C_4^{\hat{\mathbf{x}}}$ and set

$$\mathcal{R}:=E\left[\exp\left(\frac{\sqrt{-1}un}{N}\zeta_1^h\right)\right]-\left\{1-\frac{1}{2}\frac{u^2n^2}{N^2}E\left[\left(\zeta_1^h\right)^2\right]\right\}.$$

■

Remark 5.1 ***(i).*** *In the assertion of Theorem 5.1, we can freely choose the constant C. Therefore we have that if C is small (wrt $C_1^{\hat{\mathbf{x}}}$), then the bias becomes small.*

(ii). *This theorem also gives an idea on how to choose h once n or N is fixed.*

(iii). *The constants $C_3^{\hat{\mathbf{x}}}$ and $C_1^{\hat{\mathbf{x}}}$ have explicit expressions but they seem cumbersome to compute for each model. One alternative way to compute these constants is to perform a pilot simulation and estimate the constants through a histogram of the left hand side of (5.1) or (5.2).*

(iv). *This theorem can be applied to obtain the values of the constants $C_3^{\hat{\mathbf{x}}}$ and $C_1^{\hat{\mathbf{x}}}$ which later can be used to choose an appropiate value for h.*

6 Application of the Malliavin-Thalmaier formula to Finance

In this section, we compute Greeks using the Malliavin-Thalmaier Formula. The set-up of this section is rather general and does not refer to the financial issues. We refer to the reader to Fournié et. al. [5] for more details about the financial background.

We consider a random vector $F^\mu=(F_1^\mu,...,F_d^\mu)$, $\mu\in\mathbb{R}^m$; $m\in\mathbb{N}$ which depends on a parameter μ. We suppose through this section that F^μ is a.s. differentiable with respect to μ. Furthermore, we assume that $F^\mu\in(\mathbb{D}^\infty)^d$ is a nondegenerate random vector. And let $f(x_1,...,x_d)$ be a payoff function in the following class $\mathcal{A}$; [1]

$$\mathcal{A}:=\left\{f:\mathbb{R}^d\to\mathbb{R}:\begin{array}{l}\text{continuous a.e. w.r.t. Lebesgue measure,}\\ \text{and there exist constants } c,a \text{ such that } |f(\mathbf{x})|\leq\frac{c}{(1+|\mathbf{x}|)^a}\ (a>1).\end{array}\right\}.$$

[1] Note that in the case of a put option, clearly $(K-x)_+\in\mathcal{A}$. Also in digital put option case, $\mathbf{1}_{[0,K]}(x)\in\mathcal{A}$. In the call cases, the results in this section if we apply the put-call parity before calculating the Greeks.

Finally if we want to compute a Greeks for call option case $(x-K)_+$, then one uses directly g_i and g_i^h after taking the derivative. Although it is known that then a localization is needed.

Note that functions in $\mathcal{A}$ are bounded.

A greek is defined for $f \in \mathcal{A}$, as the following quantity for some $j \in \{1, ..., , m\}$

$$\frac{\partial}{\partial \mu_j} E\Big[f(F_1^\mu, ..., F_d^\mu)\Big].$$

As the study of the second derivative is similar we concentrate on the above quantity and just quote the result for second derivatives in the next section.

First we give some lemmas. For $i = 1, ..., d$ and $f \in \mathcal{A}$, set

$$g_i(\mathbf{y}) := \int_{\mathbb{R}^d} f(\mathbf{x}) \frac{\partial}{\partial x_i} Q_d(\mathbf{y} - \mathbf{x}) d\mathbf{x},$$
$$g_i^h(\mathbf{y}) := \int_{\mathbb{R}^d} f(\mathbf{x}) \frac{\partial}{\partial x_i} Q_d^h(\mathbf{y} - \mathbf{x}) d\mathbf{x}.$$

Note that $g_i^h \in C^\infty(\mathbb{R}^d)$ for $i = 1, ..., d$.

Lemma 6.1 *For $f \in \mathcal{A} \cap L^p(\mathbb{R}^d)$ $(p > 1)$ and $i = 1, ..., d$*

$$g_i^h(\mathbf{y}) \longrightarrow g_i(\mathbf{y}) \quad a.e.$$

Proof. For $\delta > 0$,

$$\int_{\mathbb{R}^d} f(\hat{\mathbf{x}}) \left\{ \frac{\partial Q_d}{\partial x_i}(\mathbf{y} - \hat{\mathbf{x}}) - \frac{\partial Q_d^h}{\partial x_i}(\mathbf{y} - \hat{\mathbf{x}}) \right\} d\hat{\mathbf{x}}$$
$$= \int_{\mathbb{R}^d} f(\hat{\mathbf{x}}) \left\{ \frac{\partial Q_d}{\partial x_i}(\mathbf{y} - \hat{\mathbf{x}}) - \frac{\partial Q_d^h}{\partial x_i}(\mathbf{y} - \hat{\mathbf{x}}) \right\} d\hat{\mathbf{x}} \tag{6.1}$$
$$+ \int_{\mathbb{R}^d} f(\hat{\mathbf{x}}) \left\{ \frac{\partial Q_d}{\partial x_i}(\mathbf{y} - \hat{\mathbf{x}}) - \frac{\partial Q_d^h}{\partial x_i}(\mathbf{y} - \hat{\mathbf{x}}) \right\} d\hat{\mathbf{x}}. \tag{6.2}$$

Note that $f \in \mathcal{A} \Rightarrow f \in L^p(\mathbb{R}^d)$ $(p > d/a)$. Then we take $\frac{d}{a} < p < d$ and $\frac{1}{p} + \frac{1}{q} = 1$. By the dominated convergence theorem, we have that for any $\delta > 0$,

$$\Big| \text{ (6.2) } \Big| \le \|f\|_p \left\| \frac{\partial Q_d}{\partial x_i}(\mathbf{y} - \cdot) - \frac{\partial Q_d^h}{\partial x_i}(\mathbf{y} - \cdot) \right\|_{q, B(\mathbf{y};\delta)^c} \longrightarrow 0, \quad (h \to 0),$$

where $\| \cdot \|_p$ denotes the $L^p(\mathbb{R}^d)$-norm and $\| \cdot \|_{q,B(\mathbf{y};\delta)^c}$ denotes the $L^q(B(\mathbf{y};\delta)^c)$-norm, $B(\mathbf{y};\delta)^c$ denotes the complement of the d-dimensional sphere with center $\mathbf{y} \in \mathbb{R}^d$ and radius $\delta > 0$.

Next we consider (6.1).

$$(6.1) = \int_{|\mathbf{y}-\hat{\mathbf{x}}| \le \delta} (f(\hat{\mathbf{x}}) - f(\mathbf{y})) \left\{ \frac{\partial}{\partial x_i} Q_d(\mathbf{y} - \hat{\mathbf{x}}) - \frac{\partial}{\partial x_i} Q_d^h(\mathbf{y} - \hat{\mathbf{x}}) \right\} d\hat{\mathbf{x}} \tag{6.3}$$
$$+ f(\mathbf{y}) \int_{|\mathbf{y}-\hat{\mathbf{x}}| \le \delta} \left\{ \frac{\partial}{\partial x_i} Q_d(\mathbf{y} - \hat{\mathbf{x}}) - \frac{\partial}{\partial x_i} Q_d^h(\mathbf{y} - \hat{\mathbf{x}}) \right\} d\hat{\mathbf{x}}.$$

As in the proof of Lemma 8.2, the second term equals 0. Therefore as $\delta \to 0$, (6.3) converges to 0 due to the continuity of f a.e. and that $\int_{|\mathbf{y}-\hat{\mathbf{x}}| \le \delta} \left| \frac{\partial}{\partial x_i} Q_d(\mathbf{y} - \hat{\mathbf{x}}) - \frac{\partial}{\partial x_i} Q_d^h(\mathbf{y} - \hat{\mathbf{x}}) \right| d\hat{\mathbf{x}} < \infty$. Therefore the result follows. ■

Remark:

From this remark and Lemma 6.1 we obtain the following convergence.

Lemma 6.2 *For $f \in \mathcal{A}$ and $i = 1, ..., d$,*

$$E\Big[g_i^h(F^\mu)\Big] \longrightarrow E\Big[g_i(F^\mu)\Big] \quad \text{as } h \to 0.$$

■

We denote the integration with respect to $p^h_{F^\mu,1}(\mathbf{x})$ by $E^h[\cdot]$. That is,

$$E^h\Big[f(F^\mu)\Big] := \int_{\mathbb{R}^d} f(\hat{\mathbf{x}})p^h_{F^\mu,1}(\hat{\mathbf{x}})d\hat{\mathbf{x}}.$$

Lemma 6.3 *For $f \in \mathcal{A}$,*

$$E\Big[f(F^\mu)\Big] = \sum_{i=1}^d E\Big[g_i(F^\mu)H_{(i)}(F^\mu;1)\Big],$$

$$E^h\Big[f(F^\mu)\Big] = \sum_{i=1}^d E\Big[g_i^h(F^\mu)H_{(i)}(F^\mu;1)\Big].$$

The proof of this lemma is straightforward. In fact, for the proof use the Malliavin-Thalmaier formula (2.4), multiply it by $f(\hat{\mathbf{x}})$ integrate and finally apply Fubini's Theorem.

Now we consider an expression of a first derivative.

Proposition 6.1 *Let $k \in \{1,...,m\}$ be fixed. Let F^μ be a nondegenerate random vector which is a.s. differentiable with respect to μ_k. Suppose that for every $i = 1,...,d$, $H_{(1,...,d,i)}(F^\mu;1)$ is a.s. differentiable in μ_k, $\frac{\partial}{\partial\mu_k}H_{(1,...,d,i)}(F^\mu;1) \in L^2(\Omega)$, and also $\frac{\partial F_j^\mu}{\partial\mu_k} \in L^2(\Omega)$ for all $j = 1,...,d$. Then we have*

$$\frac{\partial}{\partial\mu_k}E^h\Big[f(F^\mu)\Big] = \sum_{i=1}^d \frac{\partial}{\partial\mu_k}E\Big[g_i^h(F^\mu)H_{(i)}(F^\mu;1)\Big] \longrightarrow \sum_{i=1}^d \frac{\partial}{\partial\mu_k}E\Big[g_i(F^\mu)H_{(i)}(F^\mu;1)\Big] = \frac{\partial}{\partial\mu_k}E\Big[f(F^\mu)\Big].$$

Proof. Using the IBP formula d times, for $i = 1,...,d$, we have that the following equality is satisfied for $f = g_i^h$, g_i

$$E\Big[f(F^\mu)H_{(i)}\Big(F^\mu;1\Big)\Big] = E\left[\int_0^{F_d^\mu}\cdots\int_0^{F_1^\mu} f(\mathbf{z})d\mathbf{z}H_{(1,...,d,i)}\Big(F^\mu;1\Big)\right].$$

For $i = 1,...,d$, define

$$G_i^h(\mathbf{y}) := \int_0^{y_d}\cdots\int_0^{y_1} g_i^h(\mathbf{z})d\mathbf{z} \quad \text{and} \quad G_i(\mathbf{y}) := \int_0^{y_d}\cdots\int_0^{y_1} g_i(\mathbf{z})d\mathbf{z}.$$

From Lemma 8.7, we have that g_i^h, $i = 1,...,d$, has at most polynomial growth. Therefore the same property is satisfied by G_i^h for $i = 1,...,d$, and then

$$\begin{aligned}
&\frac{\partial}{\partial\mu_k}E\left[G_i^h(F^\mu)H_{(1,...,d,i)}\Big(F^\mu;1\Big)\right] \\
&= E\left[\sum_{j=1}^d \frac{\partial}{\partial y_j}G_i^h(F^\mu)\frac{\partial F_j^\mu}{\partial\mu_k}H_{(1,...,d,i)}\Big(F^\mu;1\Big)\right] + E\left[G_i^h(F^\mu)\frac{\partial}{\partial\mu_k}H_{(1,...,d,i)}\Big(F^\mu;1\Big)\right].
\end{aligned}$$

We consider the first term. Let $\mathbf{y}$ be fixed. From Lemma 8.7, g_i^h, $i = 1,...,d$, has at most polynomial growth, then it is bounded on $[0,y_1]\times\cdots\times[0,y_d]$. Hence for $j = 1,...,d$,

$$\frac{\partial}{\partial y_j}G_i^h(\mathbf{y}) \longrightarrow \frac{\partial}{\partial y_j}G_i(\mathbf{y}) \quad \text{as } h \to 0.$$

And since g_i^h, $i = 1,...,d$, has at most polynomial growth, $\frac{\partial}{\partial y_j}G_i^h$, $i,j = 1,...,d$, has also polynomial growth where the growth rate is independent of h. Hence for $i,j = 1,...,d$,

$$E\left[\frac{\partial}{\partial y_j}G_i^h(F^\mu)\frac{\partial F_j^\mu}{\partial\mu_k}H_{(1,...,d,i)}\Big(F^\mu;1\Big)\right] \longrightarrow E\left[\frac{\partial}{\partial y_j}G_i(F^\mu)\frac{\partial F_j^\mu}{\partial\mu_k}H_{(1,...,d,i)}\Big(F^\mu;1\Big)\right] \quad \text{as } h \to 0.$$

Similarly we prove the convergence of the second term; for $i = 1, ..., d$,

$$E\left[G_i^h(F^\mu)\frac{\partial}{\partial \mu_k}H_{(1,...,d,i)}\Big(F^\mu;1\Big)\right] \longrightarrow E\left[G_i(F^\mu)\frac{\partial}{\partial \mu_k}H_{(1,...,d,i)}\Big(F^\mu;1\Big)\right].$$

From here the result follows in a straightforward manner. ■

For $i, j = 1, ..., d$, define

$$g_{i,j}^h(\mathbf{y}) := \frac{\partial g_i^h}{\partial y_j}(\mathbf{y}) = \frac{\partial}{\partial y_j}\int_{\mathbb{R}^d} f(\hat{\mathbf{x}})\frac{\partial}{\partial x_i}Q_d^h(\mathbf{y}-\hat{\mathbf{x}})d\hat{\mathbf{x}}, \quad \mathbf{y} \in \mathbb{R}^d. \tag{6.4}$$

Remark 6.1 *Note that if $f \in \mathcal{A}$ then $g_{i,j}^h$ exists and is finite for $i, j = 1, ..., d$ and $\mathbf{y} \in \mathbb{R}^d$.*

Theorem 6.1 *Let $k \in \{1, ..., m\}$ be fixed and let $f \in \mathcal{A}$. Moreover, let F^μ be a nondegenerate random vector, which is a.s. differentiable with respect to μ_k. Suppose that for $j = 1, ..., d$, $\frac{\partial F_j^\mu}{\partial \mu_k} \in \mathbb{D}^\infty$. Then for $i = 1, ..., d$,*

$$\frac{\partial}{\partial \mu_k}E^h\Big[f(F^\mu)\Big] = \frac{\partial}{\partial \mu_k}\sum_{i=1}^d E\Big[g_i^h(F^\mu)H_{(i)}\Big(F^\mu;1\Big)\Big] = \sum_{i,j=1}^d E\left[g_{i,j}^h(F^\mu)H_{(i)}\left(F^\mu;\frac{\partial F_j^\mu}{\partial \mu_k}\right)\right].$$

Moreover if we assume that for all $i, j = 1, ..., d$, there exists a function $g_{i,j}$ such that $g_{i,j}^h(F^\mu) \to g_{i,j}(F^\mu)$ in $L^{1+\varepsilon}(\Omega)$ as $h \to 0$ for some $\varepsilon > 0$, then

$$\sum_{i,j=1}^d E\left[g_{i,j}(F^\mu)H_{(i)}\left(F^\mu;\frac{\partial F_j^\mu}{\partial \mu_k}\right)\right] = \frac{\partial}{\partial \mu_k}E\Big[f(F^\mu)\Big]. \tag{6.5}$$

Proof. We prove the first part by using the IBP formula. For $i = 1, ..., d$,

$$\begin{aligned}\frac{\partial}{\partial \mu_k}E\Big[g_i^h(F^\mu)H_{(i)}\Big(F^\mu;1\Big)\Big] &= \frac{\partial}{\partial \mu_k}E\left[\frac{\partial}{\partial y_i}g_i^h(F^\mu)\right] \\ &= E\left[\sum_{j=1}^d \frac{\partial^2}{\partial y_j \partial y_i}g_i^h(F^\mu)\frac{\partial F_j^\mu}{\partial \mu_k}\right] \\ &= \sum_{j=1}^d E\left[\frac{\partial}{\partial y_j}g_i^h(F^\mu)H_{(i)}\left(F^\mu;\frac{\partial F_j^\mu}{\partial \mu_k}\right)\right].\end{aligned}$$

where we have used Lemma 8.15. Therefore we obtain the first assertion. The second claim follows by taking limits. ■

Remark 6.2 ***(i).*** *Note that the expression in Theorem 6.1 is obviously not unique. In fact, we also have*

$$\frac{\partial}{\partial \mu_k}\sum_{i=1}^d E\Big[g_i^h(F^\mu)H_{(i)}\Big(F^\mu;1\Big)\Big] = \sum_{i,j=1}^d E\left[g_{i,i}^h(F^\mu)H_{(j)}\left(F^\mu;\frac{\partial F_j^\mu}{\partial \mu_k}\right)\right].$$

(ii). *In the digital put case, we have an explicit expression of $g_{i,j}$, $i, j = 1, ..., d$. That is, let $d = 2$ and $f(x_1, x_2) = \mathbf{1}(0 \le x_1 \le K_1)\mathbf{1}(0 \le x_2 \le K_2) \in \mathcal{A}$ where K_1 and K_2 are positive constants.*

$$\begin{aligned} g_{1,1}(\mathbf{y}) &= A_2\left\{\arctan\frac{y_2}{y_1} - \arctan\frac{y_2-K_2}{y_1} - \arctan\frac{y_2}{y_1-K_1} + \arctan\frac{y_2-K_2}{y_1-K_1}\right\}, \\ g_{2,2}(\mathbf{y}) &= A_2\left\{\arctan\frac{y_1}{y_2} - \arctan\frac{y_1-K_1}{y_2} - \arctan\frac{y_1}{y_2-K_2} + \arctan\frac{y_1-K_1}{y_2-K_2}\right\}.\end{aligned}$$

$$g_{1,2}(\mathbf{y}) = g_{2,1}(\mathbf{y}) = \frac{A_2}{2}\ln\left(\frac{\left(y_1^2+y_2^2\right)\left((y_1-K_1)^2+(y_2-K_2)^2\right)}{\left((y_1-K_1)^2+y_2^2\right)\left(y_1^2+(y_2-K_2)^2\right)}\right) \tag{6.6}$$

These expressions are obtained after taking limits of $g_{i,j}^h(\mathbf{y})$ as $h \to 0$ for $i, j = 1, 2$.

(iii). *In general, if $g^h_{i,j}$, $i,j=1,...,d$ has an explicit representation, then one can calculate Greeks easily. If we do not have an explicit expression for the multiple integral then one can use a suitable approximation for multiple Lebesgue integrals.*

(iv). *The case of second derivatives follows along a similar pattern and we only quote briefly the result. For more details, see [11]. Let $k,l \in \{1,...,m\}$ be fixed. Suppose that for $i=1,...,d$, $l,k=1,...,n$, $\frac{\partial F^\mu_i}{\partial \mu_k}$, $\frac{\partial F^\mu_i}{\partial \mu_l}$, $\frac{\partial^2 F^\mu_i}{\partial \mu_k \partial \mu_l} \in \mathbb{D}^\infty$. Furthermore, assume that for all $i,j=1,...,d$, there exists functions $g_{i,i}$, $g_{i,i,j}$ such that $\left(g^h_{i,i}, g^h_{i,i,j}\right)(F^\mu) \to (g_{i,i}, g_{i,i,j})(F^\mu)$ in $L^{1+\varepsilon}(\Omega)$ as $h \to 0$ for some $\varepsilon > 0$. Then we have*

$$\frac{\partial^2}{\partial \mu_l \partial \mu_k} E^h\Big[f(F^\mu)\Big] = \sum_{i,j_1=1}^{d} \left\{ \sum_{j_2=1}^{d} E\left[g^h_{i,i,j_2}(F^\mu) H_{(j_1)} \left(F^\mu; \frac{\partial F^\mu_{j_2}}{\partial \mu_l} \frac{\partial F^\mu_{j_1}}{\partial \mu_k} \right) \right] + E\left[g^h_{i,i}(F^\mu) H_{(j_1)} \left(F^\mu; \frac{\partial^2 F^\mu_{j_1}}{\partial \mu_l \partial \mu_k} \right) \right] \right\}$$

$$\longrightarrow \sum_{i,j_1=1}^{d} \left\{ \sum_{j_2=1}^{d} E\left[g_{i,i,j_2}(F^\mu) H_{(j_1)} \left(F^\mu; \frac{\partial F^\mu_{j_2}}{\partial \mu_l} \frac{\partial F^\mu_{j_1}}{\partial \mu_k} \right) \right] + E\left[g_{i,i}(F^\mu) H_{(j_1)} \left(F^\mu; \frac{\partial^2 F^\mu_{j_1}}{\partial \mu_l \partial \mu_k} \right) \right] \right\} = \frac{\partial^2}{\partial \mu_l \partial \mu_k} E\Big[f(\ldots$$

(v). *Note that we have written the approximation of the second derivative as part of the statement. This is because in some particular situations it may be more convenient to use the approximation to the second derivative rather than the limit expression itself. For example, this is the case for $d=2$ and when the second derivative coincides with the density function.*

7 Examples and Simulations

In this section, we provide some simple examples of application in two cases. In the first, we approximate the multi-dimensional log-normal density. We take this as a toy-example, to show the performance of the Malliavin-Thalmaier method with and without regularization parameter (i.e. $h > 0$ or $h = 0$). In order to be concise, we only describe the results and comment on the important issues in the toy-example. The case of the bivariate density of the Heston model is solved using the technique presented in this paper and using a finite difference scheme.

7.1 The multivariate Geometric Brownian motion

Consider the solution of the following stochastic differential equation,

$$\frac{dX^i_t}{X^i_t} = \mu_i dt + \sum_{j=1}^{d} \sigma_{ij} dW^j_t, \quad X^i_0 = x_i. \tag{7.1}$$

where $W_t = (W^1_t, ..., W^d_t)$ is a standard d-dimensional Brownian motion, μ_i and σ_{ij} are constants.

The density of $X_t = (X^1_t, ..., X^d_t)$ is the multivariate lognormal distribution. As the goal is to compare the theoretical density with the Malliavin-Thalmaier approach with and without regularization parameter h we only need to the formula (3.1) explicitly. In particular, we derive an expression for the weight $H_{(i)}(F;G)$. First, define, for $i = 1, ..., d$,

$$f_i(x) := y^0_i \exp\left(\left(\mu_i - \frac{\sum_{j=1}^{d} \sigma^2_{ij}}{2} \right) T + x \right),$$

$$X^i_T := f_i\Big(\sum_{j=1}^{d} \sigma_{ij} W^j_T\Big).$$

Then, we have using the chain rule for Malliavin derivatives that

$$DX_T^i = \left\{ \frac{\partial}{\partial x} f_i(\sum_{j=1}^d \sigma_{ij} W_T^j) \right\} D \left(\sum_{j=1}^d \sigma_{ij} W_T^j \right) = X_T^i \begin{pmatrix} \sigma_{i1} \mathbf{1}(\cdot \leq T) \\ \vdots \\ \sigma_{id} \mathbf{1}(\cdot \leq T) \end{pmatrix} .$$

Lemma 7.1 *Let F be a nondegenerate random vector then the density of $F = X_T$, solution of equation (7.1), can be expressed as*

$$p_{F,1}(\hat{\mathbf{x}}) = A_d \sum_{i=1}^d E \left[\frac{F_i - \hat{x}_i}{|F - \hat{\mathbf{x}}|^d} \sum_{j=1}^d (-1)^{i+j} \frac{\det(\Sigma_i^j)}{\det(\Sigma)} \left\{ \frac{W_T^j}{F_i} + \frac{\sigma_{ij} T}{F_i} \right\} \right] . \tag{7.2}$$

Here, $\Sigma := (\sigma_{ij})_{i,j=1,\dots,d}$ and Σ_j^i, j, $i = 1, \dots, d$, is a $(d-1) \times (d-1)$ matrix obtained from Σ by deleting row i and column j.

Proof. The expression of $H_{(i)}$ does not depend on the function f. Therefore from Proposition 2.1 and Lemma 8.16 in the Appendix, we have an representation of the density $p_{F,G}$, where $F = X_T$. ∎

Our approximation to the density is given by

$$p_{F,1}^h(\hat{\mathbf{x}}) = A_d \sum_{i=1}^d E \left[\frac{F_i - \hat{x}_i}{|F - \hat{\mathbf{x}}|_h^d} \sum_{j=1}^d (-1)^{i+j} \frac{\det(\Sigma_i^j)}{\det(\Sigma)} \left\{ \frac{W_T^j}{F_i} + \frac{\sigma_{ij} T}{F_i} \right\} \right] . \tag{7.3}$$

Now we briefly comment on the simulation results. We realized the Monte Carlo simulation of both formulas (7.3) and (7.2) and compared them with the theoretical result in the $d = 2$ dimensional case. In particular the parameters were

$$\mu = (0.01, 0.02)$$
$$\sigma = \begin{pmatrix} 0.1 & 0.2 \\ 0.3 & 0.2 \end{pmatrix} .$$

The densities were then compared. In particular the Malliavin-Thalmaier formula is highly biased in comparison with the approximative formula (7.3). The Malliavin-Thalmaier formula exibited peaks which are due to the unstable behaviour of $\frac{\partial}{\partial x_i} Q_d$. This unstability can also be observed at a local level. In comparison the regularized version behaves smoothly. The choice of $h = 0.01$ was an adhoc choice. In fact, the central limit theorem Theorem 5.1 states that the weak convergence occurs as $h \to 0$. Nevertheless one may also want to minimize the asymptotic L^2-error as it is usually done in kernel density estimation theory. This requires a minimization procedure that can be done when the constants in the formulas appearing in Sections 3 and 4 are known. In fact, they can be obtained in practice through a pilot simulation that gives the histogram of the error sequences in the central limit theorem 5.1.

7.2 Example: The Heston model

Now we consider the simulation of the joint density of the underlying price and the volatility in the Heston model. First we define Heston model [6] as follows;

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{v_t} S_t \left\{ \rho dW_t^{(2)} + \sqrt{1-\rho^2} dW_t^{(1)} \right\} , \\ dv_t &= \gamma (\theta - v_t) dt + \kappa \sqrt{v_t} dW_t^{(2)} , \end{aligned} \tag{7.4}$$

where $\mu, \gamma, \theta, \kappa$ are positive constants satisfying $\gamma\theta \geq \frac{3\kappa^2}{4}$. This condition assures that v satisfies the necessary differentiability and integrability properties and that it is strictly positive a.s. For more information, we refer the reader to Alos, Ewald [1]and Section 6.2.2. in Lamberton, Lapeyre [7]. Next we consider the following change of variables. Set $X_t := \ln(S_t/S_0) - \mu t$, $u_t := av_t$ for a positive constant a. Then

$$\begin{aligned} X_t &= X_0 - \frac{1}{2a}\int_0^t u_r dr + \frac{\rho}{\sqrt{a}}\int_0^t \sqrt{u_r} dW_r^{(2)} + \sqrt{\frac{1-\rho^2}{a}}\int_0^t \sqrt{u_r} dW_r^{(1)}, \\ u_t &= u_0 - \gamma\int_0^t u_r dr + a\gamma\theta t + \sqrt{a}\kappa\int_0^t \sqrt{u_r} dW_r^{(2)}. \end{aligned} \tag{7.5}$$

In this setting we have the following facts

(i).

$$D_s^{(2)} u_t = \kappa\sqrt{au_s}\frac{e(t)}{e(s)}\mathbf{1}_{[0,t]}(s), \tag{7.6}$$

where

$$e(t) := \exp\left(-\gamma t - \frac{a\kappa^2}{8}\int_0^t \frac{1}{u_r} dr + \frac{\sqrt{a}\kappa}{2}\int_0^t \frac{1}{\sqrt{u_r}} dW_r^{(2)}\right),$$

and

$$D_s^{(2)} e(t) = e(t)\left\{\frac{\sqrt{a}\kappa}{2\sqrt{u_s}} + \frac{a\kappa^2}{8}\int_s^t \frac{D_s^{(2)} u_r}{u_r^2} dr - \frac{\sqrt{a}\kappa}{4}\int_s^t \frac{D_s^{(2)} u_r}{u_r^{\frac{3}{2}}} dW_r^{(2)}\right\}\mathbf{1}_{[0,t]}(s).$$

Also note that $D_s^{(1)} u_t \equiv 0$ and $D_s^{(1)} e(t) \equiv 0$.

(ii).

$$D_s^{(1)} X_t = \sqrt{\frac{1-\rho^2}{a}}\sqrt{u_s}\mathbf{1}_{[0,t]}(s).$$

(iii).

$$D_s^{(2)} X_t = \left(\frac{\rho\sqrt{u_s}}{\sqrt{a}} - \frac{\kappa\sqrt{u_s}}{2\sqrt{a}e(s)}\int_s^t e(r)dr + \frac{\rho\kappa\sqrt{u_s}}{2e(s)}\int_s^t \frac{e(r)}{\sqrt{u_r}} dW_r^{(2)} + \frac{\sqrt{1-\rho^2}\kappa\sqrt{u_s}}{2e(s)}\int_s^t \frac{e(r)}{\sqrt{u_r}} dW_r^{(1)}\right)\mathbf{1}_{[0,t]}(s).$$

(iv).

$$D_w^{(2)} D_s^{(1)} X_t = \frac{\sqrt{1-\rho^2}}{2\sqrt{au_s}} D_w^{(2)} u_s \mathbf{1}_{[0,t]}(s)\mathbf{1}_{[0,s]}(w).$$

(v).

$$\begin{aligned} D_w^{(2)} D_s^{(2)} u_t &= \kappa\sqrt{au_s}\frac{e(t)}{e(s)}\left\{\frac{\kappa^3\sqrt{au_w}}{8e(w)}\int_{s\vee w}^t \frac{e(r)}{u_r^2} dr - \frac{a\kappa^2}{4}\frac{\sqrt{u_w}}{e(w)}\int_{s\vee w}^t \frac{e(r)}{u_r^{\frac{3}{2}}} dW_r^{(2)} \right. \\ &\quad \left. + \frac{\sqrt{a}\kappa}{2\sqrt{u_w}}\mathbf{1}(0 \leq s \leq w \leq t) + \frac{D_w^{(2)} u_s}{2u_s}\mathbf{1}(0 \leq w < s \leq t)\right\}. \end{aligned}$$

(vi). Calculation of $H_{(1)}(F; 1)$ for $F = (X_t, u_t)$. With the previous calculations we can apply the Bismut-Elworthy formula (see Exercise 2.3.5 in [9]), to obtain

$$H_{(1)}(F; 1) = \frac{\sqrt{a}}{\sqrt{1-\rho^2}t}\int_0^t \frac{1}{\sqrt{u_s}} dW_s^{(1)}.$$

(vii). Similarly,

$$H_{(2)}(F;1) = \frac{1}{t}\left\{\delta^{(2)}\left(\frac{1}{D_{\cdot}^{(2)}u_t}\right) - \delta^{(1)}\left(\frac{D_{\cdot}^{(2)}X_t}{D_{\cdot}^{(1)}X_t}\frac{1}{D_{\cdot}^{(2)}u_t}\right)\right\}. \tag{7.7}$$

where

$$\begin{aligned}&\int_0^t \frac{1}{D_s^{(2)}u_t}dW_s^{(2)}\\&= \frac{1}{\sqrt{a}\kappa e(t)}\int_0^t \frac{e(s)}{\sqrt{u_s}}dW_s^{(2)} + \frac{1}{2e(t)}\int_0^t \frac{e(s)}{u_s}ds + \frac{a\kappa^2}{8e(t)}\int_0^t r\frac{e(r)}{u_r^2}dr - \frac{\sqrt{a}\kappa}{4e(t)}\int_0^t r\frac{e(r)}{u_r^{\frac{3}{2}}}dW_r^{(2)},\end{aligned}$$

and

$$\begin{aligned}&\delta^{(1)}\left(\frac{D_{\cdot}^{(2)}X_t}{D_{\cdot}^{(1)}X_t}\frac{1}{D_{\cdot}^{(2)}u_t}\right)\\&= \frac{\rho}{\kappa\sqrt{a(1-\rho^2)}e(t)}\int_0^t \frac{e(s)}{\sqrt{u_s}}dW_s^{(1)} - \frac{1}{2\sqrt{a(1-\rho^2)}e(t)}\int_0^t e(r)\int_0^r \frac{1}{\sqrt{u_s}}dW_s^{(1)}dr\\&\quad + \frac{\rho}{2\sqrt{1-\rho^2}e(t)}\int_0^t \frac{e(r)}{\sqrt{u_r}}\int_0^r \frac{1}{\sqrt{u_s}}dW_s^{(1)}dW_r^{(2)} + \frac{1}{2e(t)}\int_0^t \frac{e(r)}{\sqrt{u_r}}\int_0^r \frac{1}{\sqrt{u_s}}dW_s^{(1)}dW_r^{(1)},\end{aligned}$$

where we have used the fact that u_t and $e(t)$ are independent of $W_t^{(1)}$.

We compare the simulation results of the finite difference method for the associated partial differential equation and the Malliavin-Thalmaier formula with and without regularization parameter. We observe that the finite difference method is sensible to changes in the initial condition although the value stabilizes around the values $[5.2, 5.4]$. The Malliavin-Thalmaier formula without regularization also seems to converge to a similar value but there seems to be a bias in the results probably due to the high variance of the estimates. The Malliavin-Thalmaier formula with regularization exhibits a better behavior with less variance. Confidence intervals have also been computed. As before the value of h was computed by obtaining the constants $C_1^{\hat{\mathbf{x}}}$ and $C_3^{\hat{\mathbf{x}}}$ with a pilot simulation in the central limit theorem theorem 5.1. Then one minimizes the L^2 error in a fashion similar to kernel density estimation methods.

We have not compared these results with the classical formulation that follows from equation (1.1) as this will require a long calculation of triple stochastic integrals. As noted before, when the dimension of the problem increases then the dimension of the multiple stochastic integrals in (1.1) will increase while the ones using the Malliavin-Thalmaier formula will remain of order 2.

A detailed description of the simulation study will appear elsewhere.

8 Appendix

8.1 Proof of Corollary 2.1

In this section, we give a proof of Corollary 2.1.

Lemma 8.1 *For $x_i \geq 0,\ i = 1, ..., d$, the following inequalities hold;*
(i). For $d = 2$ and $1 < \kappa < 2 - \frac{2}{p},\ p > 2$ we have

$$\int_0^{x_2}\int_0^{x_1} |Q_2(\mathbf{y})|^{\frac{p}{p-1}}dy_1dy_2 \leq \frac{\pi}{2}a_2^{-\frac{p}{p-1}}\left\{\frac{p-1}{(2-\kappa)p-2} + |\mathbf{x}|^{\frac{2p-1}{p-1}}\right\}.$$

(ii). For $p > d - 1 \geq 2$,

$$\int_0^{x_d}\cdots\int_0^{x_1} |Q_d(\mathbf{y})|^{\frac{p}{p-1}}dy_1...dy_d \leq \left(\frac{\pi}{2}\right)^{d-1}a_d^{-\frac{p}{p-1}}\left\{\frac{p-1}{2p-d} + |\mathbf{x}|^{\frac{2p-d}{p-1}}\right\}.$$

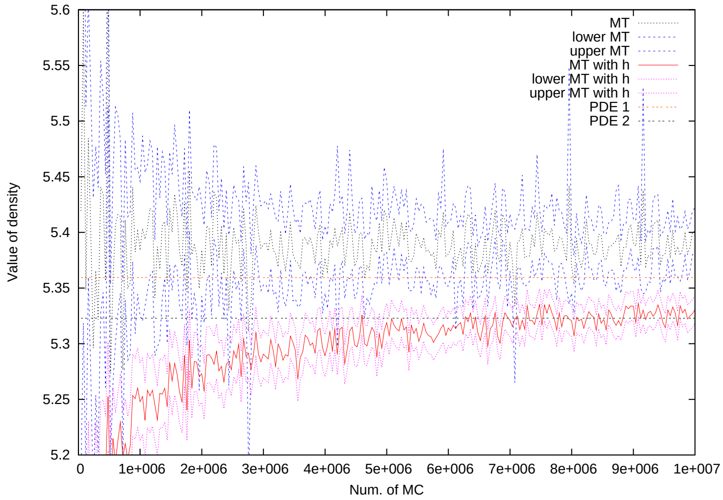

Figure 1: Simulations for the Heston model

(iii). For $p > d \geq 2$, $i = 1, ..., d$,

$$\int_0^{x_d} \cdots \int_0^{x_1} \left| \frac{\partial}{\partial y_i} Q_d(\mathbf{y}) \right|^{\frac{p}{p-1}} dy_1 ... dy_d \leq \left(\frac{\pi}{2} \right)^{d-1} A_d^{\frac{p}{p-1}} \left\{ \frac{p-1}{p-d} + |\mathbf{x}| \right\}.$$

Proof. (i). Here one performs a change of variables from rectangular to spherical coordinates and separates the region of integration in two. The first integral is bounded by the integral over the unit ball and the second on the complement. For the first one uses the inequality $\left| \ln(r) \right| < \frac{1}{r^\kappa}$ for $1 < \kappa$ and $0 < r < 1$. For the second integral, one uses that $\ln(r) < r$ for $r > 1$. Then the inequality follows after some straightforward integrations where the condition $\kappa < 2 - \frac{2}{p}$ is used. (ii) and (iii) are proved changing directly from rectangular to spherical coordinates. ■

Proof of Corollary 2.1. The goal in the proof is to show that (2.3) is satisfied.

First, note that $|Q_d(\mathbf{x})|$ and $|\frac{\partial}{\partial x_i} Q_d(\mathbf{x})|$ $(i = 1, ..., d)$ are symmetric. That is, $|Q_d(\mathbf{x})| = |Q_d(-\mathbf{x})|$ and $|\frac{\partial}{\partial x_i} Q_d(\mathbf{x})| = |\frac{\partial}{\partial x_i} Q_d(-\mathbf{x})|$.

Now, we prove that $\sup_{|\mathbf{a}| \leq R} E[|Q_2(F - \mathbf{a})|^{\frac{p}{p-1}}] < \infty$ for all $R > 0$. From the above symmetric property, the IBP formula, Lemma 8.1 and Hölder's inequality,

$$\begin{aligned}
\sup_{|\mathbf{a}| \leq R} E\left[\left| Q_2(F - \mathbf{a}) \right|^{\frac{p}{p-1}} \right] &= \sup_{|\mathbf{a}| \leq R} E\left[\frac{\partial^2}{\partial y_1 \partial y_2} \left(\int_0^{|F_2 - a_2|} \int_0^{|F_1 - a_1|} \left| Q_2(\mathbf{y}) \right|^{\frac{p}{p-1}} dy_1 dy_2 \right) \right] \\
&= \sup_{|\mathbf{a}| \leq R} E\left[\left(\int_0^{|F_2 - a_2|} \int_0^{|F_1 - a_1|} \left| Q_2(\mathbf{y}) \right|^{\frac{p}{p-1}} dy_1 dy_2 \right) H_{(1,2)}(F; 1) \right] \\
&\leq \frac{\pi}{2} a_2^{-\frac{p}{p-1}} \left\{ \frac{p-1}{(2-\kappa)p - 2} + E\left[\left(|F| + R \right)^{\frac{2p-1}{p-1} \frac{p}{p-1}} \right]^{\frac{p-1}{p}} \right\} E\left[\left| H_{(1,2)}(F; 1) \right|^p \right]^{\frac{1}{p}} < \infty.
\end{aligned}$$

Proving the inequality for general d and the derivative of Q_d follows along the same lines as above. ■

8.2 Lemmas used in the proof of Theorem 3.1

Lemma 8.2 *For $i = 1, ..., d$,*

$$A_d \left(\lim_{\varepsilon \to 0} \Phi_{i,\varepsilon}^F(\hat{\mathbf{x}}) \right) \int_0^1 \frac{(r^2 + h)^{\frac{d}{2}} - r^d}{(r^2 + h)^{\frac{d}{2}}} dr \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cdots \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \Theta_i c_1^{d-2} \cdots c_{d-2} d\theta_1 ... d\theta_{d-1} = 0.$$

Proof. We know by Lemma 3.1 that $\lim_{\varepsilon\to 0}\Phi^F_{i,\varepsilon}(\hat{\mathbf{x}})$ is finite. Similarly, for fixed h we have that $\int_0^1 \frac{(r^2+h)^{\frac{d}{2}}-r^d}{(r^2+h)^{\frac{d}{2}}}dr$ is finite. Therefore the result follows from $\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\sin\theta\cos^n\theta d\theta = 0$ for $n\in\mathbb{N}$. ∎

Lemma 8.3 *For $i,j = 1,...,d$,*

$$A_d\left(\lim_{\varepsilon\to 0}\frac{\partial}{\partial y_j}\Phi^F_{i,\varepsilon}(\hat{\mathbf{x}})\right)\int_0^1 r\frac{(r^2+h)^{\frac{d}{2}}-r^d}{(r^2+h)^{\frac{d}{2}}}dr\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\Theta_i\Theta_j c_1^{d-2}\cdots c_{d-2}d\theta_1\cdots d\theta_{d-1}$$
$$=\begin{cases} C^{\hat{\mathbf{x}}}_{1,i}h\ln\frac{1}{h}+C^{\hat{\mathbf{x}}}_{2,i}h+o(h) & ; \quad \text{for } i=j\\ 0 & ; \quad \text{for } i\neq j,\end{cases}$$

where

$$C^{\hat{\mathbf{x}}}_{1,i} := \frac{d}{4}A_d\left(\lim_{\varepsilon\to 0}\frac{\partial}{\partial y_i}\Phi^F_{i,\varepsilon}(\hat{\mathbf{x}})\right)\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\Theta_i^2 c_1^{d-2}\cdots c_{d-2}d\theta_1\cdots d\theta_{d-1},$$
$$C^{\hat{\mathbf{x}}}_{2,i} := A_d\left(\lim_{\varepsilon\to 0}\frac{\partial}{\partial y_i}\Phi^F_{i,\varepsilon}(\hat{\mathbf{x}})\right)\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\Theta_i^2 c_1^{d-2}\cdots c_{d-2}d\theta_1\cdots d\theta_{d-1}$$
$$\times\left[\int_0^1 u\frac{(u^2+1)^{\frac{d}{2}}-u^d}{(u^2+1)^{\frac{d}{2}}}du+\frac{1}{4}\ln\frac{1}{2^{d-1}+2^{\frac{d}{2}-1}}+M_d^0\right],$$

and M_d^0 is a constant (defined in the proof).

Proof. In the case of $i\neq j$, using the same argument of Lemma 8.2, the result follows. Next in the case of $i=j$, note that

$$\left|A_d\left(\lim_{\varepsilon\to 0}\frac{\partial}{\partial y_i}\Phi^F_{i,\varepsilon}(\hat{\mathbf{x}})\right)\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\Theta_i^2 c_1^{d-2}\cdots c_{d-2}d\theta_1\cdots d\theta_{d-1}\right|<\infty.$$

Set $u=\frac{r}{\sqrt{h}}$.

$$\int_0^1 r\frac{(r^2+h)^{\frac{d}{2}}-r^d}{(r^2+h)^{\frac{d}{2}}}dr = h\int_0^1 u\frac{(u^2+1)^{\frac{d}{2}}-u^d}{(u^2+1)^{\frac{d}{2}}}du + h\int_1^{\frac{1}{\sqrt{h}}}u\frac{\sum_{l=1}^d\binom{d}{l}u^{2(d-l)}}{(u^2+1)^{\frac{d}{2}}\{(u^2+1)^{\frac{d}{2}}+u^d\}}du.$$

Clearly the first term on the right hand side has finite value. Next we consider the second term.

$$h\int_1^{\frac{1}{\sqrt{h}}}u\frac{\sum_{l=1}^d\binom{d}{l}u^{2(d-l)}}{(u^2+1)^{\frac{d}{2}}\{(u^2+1)^{\frac{d}{2}}+u^d\}}du$$
$$=\frac{h}{4}\int_1^{\frac{1}{\sqrt{h}}}\frac{2du(u^2+1)^{d-1}+\frac{d}{2}(4u^3+2u)(u^4+u^2)^{\frac{d}{2}-1}}{(u^2+1)^d+(u^4+u^2)^{\frac{d}{2}}}du$$
$$+h\left[-\frac{d}{4}\int_1^{\frac{1}{\sqrt{h}}}\left(2u\sum_{l=1}^{d-1}\binom{d-1}{l}u^{2(d-1-l)}+2u^{2d-1}\left(\left(1+\frac{1}{2u^2}\right)\left(1+\frac{1}{u^2}\right)^{\frac{d}{2}-1}-1\right)\right)\right.$$
$$\left.\times\left((u^2+1)^d+(u^4+u^2)^{\frac{d}{2}}\right)^{-1}du+\int_1^{\frac{1}{\sqrt{h}}}\frac{\sum_{l=2}^d\binom{d}{l}u^{2(d-l)}}{(u^2+1)^d+(u^4+u^2)^{\frac{d}{2}}}du\right]$$
$$=\frac{h}{4}\left\{\ln\left((1+h)^d+(1+h)^{\frac{d}{2}}\right)+d\ln\frac{1}{h}-\ln\left(2^d+2^{\frac{d}{2}}\right)\right\}+hM_d^h,$$

where M_d^h is defined by the term within brackets [ ]. Note that the integrands in M_d^h are of order $O(u^2)$ as $u\to\infty$ and therefore integrable. Then we define $M_d^0 := \lim_{h\to 0}M_d^h$. Hence

$$\lim_{h\to 0}\frac{\frac{h}{4}\left\{\ln\left((1+h)^d+(1+h)^{\frac{d}{2}}\right)-\ln\left(2^d+2^{\frac{d}{2}}\right)\right\}+hM_d^h}{h}=\frac{1}{4}\left(\ln(2)+\ln\left(2^d+2^{\frac{d}{2}}\right)\right)+M_d^0.$$ ∎

Lemma 8.4 *For $i, j, k = 1, ..., d$,*

$$\frac{A_d}{2}\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\int_0^1 r^2\frac{(r^2+h)^{\frac{d}{2}}-r^d}{(r^2+h)^{\frac{d}{2}}}\Theta_i\Theta_j\Theta_k c_1^{d-2}\cdots c_{d-2} \times\left(\lim_{\varepsilon\to 0}\int_0^1\frac{\partial^2}{\partial y_k\partial y_j}\Phi_{i,\varepsilon}^F(\mathbf{x}+\gamma r\Theta)d\gamma dr d\theta_1\cdots d\theta_{d-1}\right)=C_{3,i,j,k}^{\hat{\mathbf{x}}}h+o(h), \quad (8.1)$$

where

$$C_{3,i,j,k}^{\hat{\mathbf{x}}}:=\frac{dA_d}{4}\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\int_0^1\Theta_i\Theta_j\Theta_k c_1^{d-2}\cdots c_{d-2}\left(\lim_{\varepsilon\to 0}\int_0^1\frac{\partial^2}{\partial y_k\partial y_j}\Phi_{i,\varepsilon}^F(\hat{\mathbf{x}}+\gamma r\Theta)d\gamma\right)dr d\theta_1\cdots d\theta_{d-1}.$$

Proof. From l'Hôpital's Rule,

$$\lim_{h\to 0}\frac{(\text{LHS of (8.1)})}{h} = \lim_{h\to 0}\frac{A_d}{2}\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\int_0^1\frac{d}{2}\frac{r^{d+2}}{(r^2+h)^{\frac{d+2}{2}}}\Theta_i\Theta_j\Theta_k c_1^{d-2}\cdots c_{d-2} \times\left(\lim_{\varepsilon\to 0}\int_0^1\frac{\partial^2}{\partial y_k\partial y_j}\Phi_{i,\varepsilon}^F(\hat{\mathbf{x}}+\gamma r\Theta)d\gamma\right)dr d\theta_1\cdots d\theta_{d-1}.$$

The result follows after applying the bounded convergence theorem. ■

Lemma 8.5 *For i=1,...,d,*

$$A_d\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\int_1^\infty\frac{(r^2+h)^{\frac{d}{2}}-r^d}{(r^2+h)^{\frac{d}{2}}}\Theta_i c_1^{d-2}\cdots c_{d-2}\left(\lim_{\varepsilon\to 0}\Phi_{i,\varepsilon}^F(\hat{\mathbf{x}}+r\Theta)\right)dr d\theta_1...d\theta_{d-1} = C_{4,i}^{\hat{\mathbf{x}}}h+o(h), \quad (8.2)$$

where

$$C_{4,i}^{\hat{\mathbf{x}}}:=\frac{d}{2}A_d\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\int_1^\infty\frac{1}{r^2}\Theta_i c_1^{d-2}\cdots c_{d-2}\left(\lim_{\varepsilon\to 0}\Phi_{i,\varepsilon}^F(\hat{\mathbf{x}}+r\Theta)\right)dr d\theta_1...d\theta_{d-1}.$$

Proof. As in the previous Lemma, from l'Hôpital's Rule, we have

$$\lim_{h\to 0}\frac{(\text{LHS of (8.2)})}{h} = \lim_{h\to 0}A_d\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\int_1^\infty\frac{d}{2}\frac{r^d}{(r^2+h)^{\frac{d}{2}+1}}\Theta_i c_1^{d-2}\cdots c_{d-2}\left(\lim_{\varepsilon\to 0}\Phi_{i,\varepsilon}^F(\hat{\mathbf{x}}+r\Theta)\right)dr d\theta_1...d\theta_{d-1}.$$

The result follows from the dominated convergence theorem. ■

8.3 Lemmas used in the proof of Theorem 4.1

We provide some lemmas for Section 4.1. We use the same notations and assumptions in Section 4.1.

Lemma 8.6 *For $i, j = 1, 2$,*

$$A_2^2\left(\lim_{\varepsilon\to 0}\hat{\Phi}_{i,j,\varepsilon}^F(\hat{\mathbf{x}})\right)\int_0^{2|\hat{\mathbf{x}}|+1}\frac{r^3}{(r^2+h)^2}dr\int_0^{2\pi}\Theta_i\Theta_j d\theta_1=\begin{cases}C_{5,i}^{\hat{\mathbf{x}}}\ln\frac{1}{h}+O(1) & ;\quad \text{for } i=j\\ 0 & ;\quad \text{for } i\neq j,\end{cases} \quad (8.3)$$

where

$$C_{5,i}^{\hat{\mathbf{x}}}=\frac{\pi}{2}A_2^2\left(\lim_{\varepsilon\to 0}\hat{\Phi}_{i,i,\varepsilon}^F(\hat{\mathbf{x}})\right).$$

Proof. In the case of $i \neq j$, (LHS of (8.3))= 0, as in Lemma 8.2. Set $u := \frac{r}{\sqrt{h}}$. In the case of $i = j$, then

$$\text{(LHS of (8.3))} = \pi A_2^2 \left(\lim_{\varepsilon\to 0} \hat{\Phi}^F_{i,i,\varepsilon}(\hat{\mathbf{x}})\right) \left\{ \frac{1}{4}\int_0^{\frac{2|\hat{\mathbf{x}}|+1}{\sqrt{h}}} \frac{4u^3+4u}{(u^2+1)^2}du - \int_0^{\frac{2|\hat{\mathbf{x}}|+1}{\sqrt{h}}} \frac{u}{(u^2+1)^2}du \right\}.$$

The first term in the brackets can be computed as follows:

$$\frac{1}{4}\int_0^{\frac{2|\hat{\mathbf{x}}|+1}{\sqrt{h}}} \frac{4u^3+4u}{(u^2+1)^2}du = \frac{1}{2}\left\{ \ln\left((2|\hat{\mathbf{x}}|+1)^2+h\right) + \ln\frac{1}{h}\right\}.$$

We can easily find that the second term is bounded uniformly in h. ■

Lemma 8.7 *For $i, j, k = 1, 2$,*

$$\left| A_2^2 \int_0^{2\pi}\int_0^{2|\hat{\mathbf{x}}|+1} \frac{r^4\Theta_i\Theta_j\Theta_k}{(r^2+h)^2} \left\{ \lim_{\varepsilon\to 0}\int_0^1 \frac{\partial}{\partial y_k}\hat{\Phi}^F_{i,j,\varepsilon}(\hat{\mathbf{x}}+\gamma r\Theta)d\gamma \right\} dr d\theta_1 \right| \le C_6^{\hat{\mathbf{x}}},$$

where $C_6^{\hat{\mathbf{x}}}$ is a positive constant which depends on $\hat{\mathbf{x}}$.

Proof. By Lemma 3.1, $\lim_{\varepsilon\to 0}\int_0^1 \frac{\partial}{\partial y_k}\hat{\Phi}^F_{i,j,\varepsilon}(\hat{\mathbf{x}}+\gamma r\Theta)d\gamma$ is uniformly bounded. Therefore the result follows. ■

Lemma 8.8 *For $i, j = 1, 2$,*

$$\left| A_2^2 \int_0^{2\pi}\int_{2|\hat{\mathbf{x}}|+1}^{\infty} r\frac{r^2\Theta_i\Theta_j}{(r^2+h)^2} \left(\lim_{\varepsilon\to 0}\hat{\Phi}^F_{i,j,\varepsilon}(r\Theta+\hat{\mathbf{x}})\right) dr d\theta_1 \right| \le C_7^{\hat{\mathbf{x}}},$$

where $C_7^{\hat{\mathbf{x}}}$ is a positive constant which depends on $\hat{\mathbf{x}}$.

Proof. From Lemma 3.1,

$$\begin{aligned}\left| A_2^2 \int_0^{2\pi}\int_{2|\hat{\mathbf{x}}|+1}^{\infty} r\frac{r^2\Theta_i\Theta_j}{(r^2+h)^2} \left(\lim_{\varepsilon\to 0}\hat{\Phi}^F_{i,j,\varepsilon}(r\Theta+\hat{\mathbf{x}})\right) dr d\theta_1 \right| &\le A_2^2 \int_0^{2\pi}\int_{2|\hat{\mathbf{x}}|+1}^{\infty} \frac{r^3}{(r^2+h)^2}\frac{C}{1+|r\Theta+\hat{\mathbf{x}}|^2} dr d\theta_1 \\ &\le \frac{C'}{1+(|\hat{\mathbf{x}}|+1)^2},\end{aligned}$$

where C' is a positive constant. ■

8.4 Lemmas used in the proof of Theorem 4.2

We provide the lemmas used in Section 4.2. We will use the same notations and assumptions in Section 4.2.

Lemma 8.9 *Set $I(n,m) = \int \sin^n x \cos^m x dx$ for $n+m \neq 0$. Then*

$$I(n,m) = -\frac{\sin^{n-1}x\cos^{m+1}x}{n+m} + \frac{n-1}{n+m}I(n-2,m) = \frac{\sin^{n+1}x\cos^{m-1}x}{n+m} + \frac{m-1}{n+m}I(n,m-2).$$

Proof. This is proved using the IBP formula for Lebesgue integrals. ■

Lemma 8.10 *For $i,j=1,...,d$,*

$$A_d^2\left\{\lim_{\varepsilon\to 0}\hat{\Phi}^F_{i,j,\varepsilon}(\hat{\mathbf{x}})\right\}\int_0^1\frac{r^{d+1}}{(r^2+h)^d}dr\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\Theta_i\Theta_j c_1^{d-2}\cdots c_{d-2}d\theta_1...d\theta_{d-1}$$

$$=\begin{cases} C^{\hat{\mathbf{x}}}_{8,i}\dfrac{1}{h^{\frac{d}{2}-1}}+o\left(\dfrac{1}{h^{\frac{d}{2}-1}}\right) & ;\quad \text{for } i=j \\ 0 & ;\quad \text{for } i\neq j, \end{cases} \tag{8.4}$$

where

$$C^{\hat{\mathbf{x}}}_{8,i}=\begin{cases} \dfrac{3\pi}{16}A_d^2\left\{\lim_{\varepsilon\to 0}\hat{\Phi}^F_{i,i,\varepsilon}(\hat{\mathbf{x}})\right\}\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\Theta_i^2 c_1 d\theta_1 d\theta_2 & (d=3), \\ \dfrac{1}{d-2}\left(\prod_{k=0}^{\frac{d}{2}-1}\dfrac{2+2k}{d+2k}\right)A_d^2\left\{\lim_{\varepsilon\to 0}\hat{\Phi}^F_{i,i,\varepsilon}(\hat{\mathbf{x}})\right\}\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\Theta_i^2 c_1^{d-2}\cdots c_{d-2}d\theta_1...d\theta_{d-1} & (d\geq 4:\ \text{even}), \\ \dfrac{\pi}{4}\left(\prod_{k=0}^{\frac{d-7}{2}}\dfrac{3+2k}{4+2k}\right)\left(\prod_{k=0}^{\frac{d-1}{2}}\dfrac{1+2k}{d-1+2k}\right)A_d^2\left\{\lim_{\varepsilon\to 0}\hat{\Phi}^F_{i,i,\varepsilon}(\hat{\mathbf{x}})\right\}\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\Theta_i^2 c_1^{d-2}\cdots c_{d-2}d\theta_1...d\theta_{d-1} & (d\geq 5:\ \text{odd}), \end{cases}$$

where if $d=5$, then we define $\prod_{k=0}^{\frac{d-7}{2}}\frac{3+2k}{2+2k}=1$.

Proof. In the case of $i\neq j$, (LHS of (8.4))$=0$, as in Lemma 8.2.
In the case that $i=j$ we perform successively the following changes of variables $u:=\frac{r}{\sqrt{h}}$, $u=\tan\tau$ and $\nu=\arctan\frac{1}{\sqrt{h}}$ to obtain that

$$\int_0^1\frac{r^{d+1}}{(r^2+h)^d}dr=\frac{1}{h^{\frac{d}{2}-1}}\int_0^{\frac{1}{\sqrt{h}}}\frac{u^{d+1}}{(u^2+1)^d}du=\frac{1}{h^{\frac{d}{2}-1}}\int_0^{\nu}\sin^{d+1}\tau\cos^{d-3}\tau d\tau.$$

(i). In the case of $d=3$. Then

$$\frac{1}{h^{\frac{1}{2}}}\left(\int_0^{\nu}\sin^4\tau d\tau-\frac{3\pi}{16}\right)+\frac{3\pi}{16}\frac{1}{h^{\frac{1}{2}}}=\frac{3\pi}{16}\frac{1}{h^{\frac{1}{2}}}+o\left(\frac{1}{h^{\frac{1}{2}}}\right).$$

(ii). In the case that d (≥ 4) and even, we have from Lemma 8.9 that

$$\frac{1}{h^{\frac{d}{2}-1}}\left(\int_0^{\nu}\sin^{d+1}\tau\cos^{d-3}\tau d\tau-\frac{1}{d-2}\prod_{k=0}^{\frac{d}{2}-1}\frac{2+2k}{d+2k}\right)+\frac{1}{h^{\frac{d}{2}-1}}\frac{1}{d-2}\prod_{k=0}^{\frac{d}{2}-1}\frac{2+2k}{d+2k}$$
$$=\frac{1}{h^{\frac{d}{2}-1}}\frac{1}{d-2}\prod_{k=0}^{\frac{d}{2}-1}\frac{2+2k}{d+2k}+o\left(\frac{1}{h^{\frac{d}{2}-1}}\right).$$

(iii). In the case that d (≥ 5) is odd. From Lemma 8.9,

$$\frac{1}{h^{\frac{d}{2}-1}}\left(\int_0^{\nu}\sin^{d+1}\tau\cos^{d-3}\tau d\tau-\frac{\pi}{4}\left(\prod_{k=0}^{\frac{d-7}{2}}\frac{3+2k}{4+2k}\right)\left(\prod_{k=0}^{\frac{d-1}{2}}\frac{1+2k}{d-1+2k}\right)\right)$$
$$+\frac{\pi}{4}\left(\prod_{k=0}^{\frac{d-7}{2}}\frac{3+2k}{4+2k}\right)\left(\prod_{k=0}^{\frac{d-1}{2}}\frac{1+2k}{d-1+2k}\right)\frac{1}{h^{\frac{d}{2}-1}}$$
$$=\frac{\pi}{4}\left(\prod_{k=0}^{\frac{d-7}{2}}\frac{3+2k}{4+2k}\right)\left(\prod_{k=0}^{\frac{d-1}{2}}\frac{1+2k}{d-1+2k}\right)\frac{1}{h^{\frac{d}{2}-1}}+o\left(\frac{1}{h^{\frac{d}{2}-1}}\right).$$

■

Lemma 8.11 *For $i, j, k = 1, ..., d$,*

$$A_d^2 \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cdots \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^1 \frac{r^{d+2}\Theta_i\Theta_j\Theta_k}{(r^2+h)^d} c_1^{d-2} \cdots c_{d-2} \left\{ \lim_{\varepsilon\to 0} \int_0^1 \frac{\partial}{\partial y_k} \hat{\Phi}^F_{i,j,\varepsilon}(\hat{\mathbf{x}} + \gamma r\Theta) d\gamma \right\} dr d\theta_1 ... d\theta_{d-1}$$
$$= \begin{cases} O\left(\ln \frac{1}{h}\right) & ; \quad \textit{for } d = 3 \\ O\left(\frac{1}{h^{\frac{d-3}{2}}}\right) & ; \quad \textit{for } d \geq 4. \end{cases}$$

Proof. By Lemma 3.1, $\lim_{\varepsilon\to 0}(\int_0^1 \frac{\partial}{\partial y_k} \hat{\Phi}^F_{i,j,\varepsilon}(\mathbf{x} + \gamma r\Theta) d\gamma)$ is bounded. Set $u = \frac{r}{\sqrt{h}}$. Then

$$\int_0^1 \frac{r^{d+2}}{(r^2+h)^d} dr \leq \frac{1}{h^{\frac{d-3}{2}}} \int_0^1 \frac{u^{d+2}}{(u^2+1)^d} du + \frac{1}{h^{\frac{d-3}{2}}} \int_1^{\frac{1}{\sqrt{h}}} \frac{1}{u^{d-2}} du.$$

Hence the result follows. ■

Lemma 8.12 *For $i, j = 1, ..., d$, there exists some positive constant C such that*

$$\left| A_d^2 \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cdots \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_1^\infty \frac{r^{d+1}\Theta_i\Theta_j}{(r^2+h)^d} c_1^{d-2} \cdots c_{d-2} \left(\lim_{\varepsilon\to 0} \hat{\Phi}^F_{i,j,\varepsilon}(r\Theta + \hat{\mathbf{x}}) \right) dr d\theta_1 ... d\theta_{d-1} \right| \leq C.$$

Proof. By Lemma 3.1, $\lim_{\varepsilon\to 0} \hat{\Phi}^F_{i,j,\varepsilon}(r\Theta + \mathbf{x})$ is bounded. Then the result follows. ■

8.5 Lemmas used in the proof of Theorem 5.1

In this section, we give some lemmas used to prove the central limit theorem.

Lemma 8.13 *For any $d \geq 2$ and $0 < p < \frac{1}{2}$, we have $|N \times \mathcal{R}| \leq o(h^p)$.*

Proof. Generally, for $n \in \mathbb{N}$ and $x \in \mathbb{R}$, we have

$$\left| e^{\sqrt{-1}x} - \sum_{k=0}^n \frac{(\sqrt{-1}x)^k}{k!} \right| \leq \frac{|x|^{n+1}}{(n+1)!}.$$

Then it is enough to prove the following; For any $d \geq 2$, we have $E[|\zeta_1^h|^3] \leq O(1/h^{d-\frac{3}{2}})$. In fact,

$$E\left[\left|\zeta_1^h\right|^3\right] \leq E\left[\left|\sum_{i=1}^d \frac{\partial}{\partial x_i} Q_d^h(F^{(1)} - \hat{\mathbf{x}}) H_{(i)}(F; G)^{(1)}\right|^3\right] + 3E\left[\left(\sum_{i=1}^d \frac{\partial}{\partial x_i} Q_d^h(F^{(1)} - \hat{\mathbf{x}}) H_{(i)}(F; G)^{(1)}\right)^2\right] p^h_{F^{(1)},G}(\hat{\mathbf{x}})$$
$$+ 3E\left[\left|\sum_{i=1}^d \frac{\partial}{\partial x_i} Q_d^h(F^{(1)} - \hat{\mathbf{x}}) H_{(i)}(F; G)^{(1)}\right|\right] \left(p^h_{F^{(1)},G}(\hat{\mathbf{x}})\right)^2 + \left(p^h_{F^{(1)},G}(\hat{\mathbf{x}})\right)^3. \tag{8.5}$$

The second and fourth term have already been studied in Theorems 3.1, 4.1 and 4.2. Hence we estimate the first and third term.

(1). Define

$$\Phi^{i,j,k}_{F^{(1)}}(\mathbf{y}) := E\left[\left|H_{(i)}(F; G)^{(1)} H_{(j)}(F; G)^{(1)} H_{(k)}(F; G)^{(1)}\right| \ \Big| \ F^{(1)} = \mathbf{y}\right] p_{F^{(1)},1}(\mathbf{y})$$

for $i, j, k = 1, ..., d$. Using spherical coordinates and Lemma 3.1, the first term of (8.5) is first divided into two terms as follows

$$E\left[\left|\sum_{i=1}^d \frac{\partial}{\partial x_i} Q_d^h(F^{(1)} - \hat{\mathbf{x}}) H_{(i)}(F; 1)^{(1)}\right|^3\right] \tag{8.6}$$
$$\leq A_d^3 \sum_{i,j,k=1}^d \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cdots \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left(\int_0^1 + \int_1^\infty \right) \frac{r^{d+2}|\Theta_i\Theta_j\Theta_k|}{\{r^2+h\}^{\frac{3d}{2}}} \Phi^{i,j,k}_{F^{(1)}}(\hat{\mathbf{x}} + r\Theta) c_1^{d-2} \cdots c_{d-2} dr d\theta_1 ... d\theta_{d-1}.$$

We can easily check integrability of the second term of (8.6). We consider the first term. Set $r = \sqrt{h}\tan(\tau)$, $\beta := \arctan\frac{1}{\sqrt{h}}$. Then there exists a positive constant M such that

$$\Big(\text{First term of (8.6)}\Big) \leq \frac{M}{h^{d-\frac{3}{2}}}\int_0^\beta \sin^{d+2}\tau\cos^{2d-4}\tau d\tau \times \int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\sum_{i,j,k=1}^d |\Theta_i\Theta_j\Theta_k| c_1^{d-2}\cdots c_{d-2}d\theta_1...d\theta_{d-1}.$$

Then we obtain the order $\frac{1}{h^{d-\frac{3}{2}}}$.

(2). Next we calculate the third term of (8.5). Set $|\Phi^i_{F^{(1)}}(\mathbf{y})| := E[|H_{(i)}(F;G)^{(1)}| \mid F^{(1)} = \mathbf{y}\,]p_{F^{(1)},1}(\mathbf{y})$ for $i = 1, ..., d$. By using the spherical coordinates and Lemma 3.1

$$\begin{aligned}
&E\left[\left|\sum_{i=1}^d \frac{\partial}{\partial x_i}Q^h_d(F-\hat{\mathbf{x}})H_{(i)}(F;1)\right|\right] \\
&\leq \sum_{i=1}^d \int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cdots\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\left\{\int_0^1 + \int_1^\infty\right\}\frac{r|\Theta_i|}{\{r^2+h\}^{\frac{d}{2}}}|\Phi^i_F(\hat{\mathbf{x}}+r\Theta)|r^{d-1}c_1^{d-2}\cdots c_{d-2}drd\theta_1...d\theta_{d-1} \\
&\leq C_d\frac{1}{\sqrt{h}} + C'_d,
\end{aligned}$$

where C_d and C'_d are some constants. This completes the proof.

Therefore as a result of above, we have our conclusion. ■

Lemma 8.14

$$E\left[\left(\zeta^h_1\right)^2\right] = \begin{cases} C_3^{\hat{\mathbf{x}}}\ln\frac{1}{h} + O(1) & ; \quad \textit{for } d=2 \\ C_4^{\hat{\mathbf{x}}}\frac{1}{h^{\frac{d}{2}-1}} + o\left(\frac{1}{h^{\frac{d}{2}-1}}\right) & ; \quad \textit{for } d \geq 3. \end{cases}$$

Proof. In the case of $d = 2$, the result follows from Theorem 3.1 and Theorem 4.1. In the case $d \geq 3$ it follows from Theorem 3.1 and Theorem 4.2. ■

8.6 Lemma for Section 6

Lemma 8.15 *Assume that $f \in \mathcal{A}$ then we have that for $i = 1, ..., d$,*

$$|g_i(\mathbf{y})| \leq a|\mathbf{y}| + b \quad \text{and} \quad |g^h_i(\mathbf{y})| \leq a|\mathbf{y}| + b, \tag{8.7}$$

where a and b are constants which depend on d and are independent of h.

Proof. Equation (8.7) follows easily from the assumptions on f. ■

8.7 Lemma for Section 7.1

Here we obtain the weight $H_{(i)}$.

Lemma 8.16 *Let X_T, be the solution of equation (7.1), then an expression for $H_{(j)}$ is*

$$H_{(j)}(X_T;1) = \frac{1}{T}\sum_{i=1}^d(-1)^{i+j}\frac{\det(\Sigma^i_j)}{\det(\Sigma)}\left\{\frac{W^i_T}{X^j_T} + \frac{\sigma_{ji}T}{X^j_T}\right\},$$

where $j = 1, ..., d$, $\Sigma := (\sigma_{ij})_{i,j=1,...,d}$ and Σ^i_j, j, $i = 1, ..., d$, is a $(d-1)\times(d-1)$ matrix obtained from Σ by deleting row i and column j.

Proof. Let $f \in C_0^1(\mathbb{R}^d)$ and e_i be a unit vector whose ith component is 1 and the other components are 0. For $l = 1, ..., d$, we have using the chain rule

$$D^i f(X_T) = \sum_{j=1}^{d} \frac{\partial}{\partial x_j} f(X_T) D^i X_T^j = \sum_{j=1}^{d} \frac{\partial}{\partial x_j} f(X_T) X_T^i \sigma_{ij}.$$

If we consider the above as a set of equations for $i = 1, ..., d$ where the unknowns are $\frac{\partial}{\partial x_j} f(X_T)$, we can solve this set of simultaneous equation using Cramer's formula and obtain

$$E\left[\frac{\partial}{\partial x_j} f(X_T)\right] = E\left[\frac{1}{X_T^j \det(\Sigma)} \mathbf{1}(\cdot \leq T) \sum_{i=1}^{d} (-1)^{i+j} \det\left(\Sigma_j^i\right) < Df(X_T), e_i >_{\mathbb{R}^d}\right],$$

where $< \cdot, \cdot >_{\mathbb{R}^d}$ is the inner product in $\mathbb{R}^d$. Then using that the Skorohod integral, δ, is the dual operator of D we have that

$$E\left[\frac{\partial}{\partial x_j} f(X_T)\right] = E\left[\frac{1}{\det(\Sigma)} \sum_{i=1}^{d} (-1)^{i+j} \det\left(\Sigma_j^i\right) \delta^i \left(\frac{\mathbf{1}(\cdot \leq T)}{X_T^j} e_i\right) f(X_T)\right]$$

Then by (2.2), $\delta^i \left(\frac{\mathbf{1}(\cdot \leq T)}{X_T^j}\right) = \frac{W_T^i}{X_T^j} + \frac{\sigma_{ji} T}{X_T^j}$ and from here the result follows. ■

References

[1] E. Alos, C.-O. Ewald, *Malliavin differentiability of the Heston Volatility and applications to option pricing*, to appear in Advances of Applied Probability 40(1), 2008

[2] V. Bally, L. Caramellino, *Lower bounds for the density of Ito processes under weak regularity assumptions*, working paper.

[4] L.C. Evans, *Partial differential equations*, Graduate studies in Mathematics, Volume 19, American Mathematical Society, 1998.

[5] E. Fournié, J.-M. Lasry, J. Lebuchoux, P.-L. Lions, N. Touzi, *Applications of Malliavin calculus to Monte Carlo methods in finance*, Finance Stoch., 3 (4), 391-412, 1999.

[6] S.L. Heston, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, The Review of Financial Studies, Vol. 6, No. 2, 327-343, 1993.

[7] D. Lamberton, B. Lapeyre, *Introduction to stochastic calculus applied to finance*, Chapman & Hall, 1996.

[8] P. Malliavin, A. Thalmaier, *Stochastic calculus of variations in mathematical finance*, Springer Finance, Springer-Verlag, Berlin, 2006.

[9] D. Nualart, *The Malliavin calculus and related topics (Second edition)*, Probability and its Applications (New York), Springer-Verlag, Berlin, 2006.

[10] M. Sanz-Solé, *Malliavin Calculus with applications to stochastic partial differential equations*, EPFL Press, 2005.

[11] K. Yasuda, *Kernel Density Estimation: The Malliavin-Thalmaier formula and Bayesian parameter estimation.* PhD thesis. Osaka University, March 2008.